LIGNOCELLULOSIC BIOMASSES AS SOURCES OF FERMENTABLE SUGARS AND BIOCATALYSTS FOR BIOREFINERY

Salvatore Montella

Dottorato in Scienze Biotecnologiche – XXVIII ciclo
Indirizzo Biotecnologie Molecolari e Industriali
Università di Napoli Federico II

Dottorato in Scienze Biotecnologiche – XXVIII ciclo
Indirizzo Biotecnologie Molecolari e Industriali
Università di Napoli Federico II

LIGNOCELLULOSIC BIOMASSES AS SOURCES OF FERMENTABLE SUGARS AND BIOCATALYSTS FOR BIOREFINERY

Salvatore Montella

Dottorando:	Salvatore Montella
Relatore:	Prof. Vincenza Faraco
Coordinatore:	Prof. Giovanni Sannia

Alla mia famiglia

Abstract

The lignocellulosic biomasses constitute the Earth's most abundant repository of carbon, thus the future development of processes that exploit this raw materials as sources of fermentable sugars represents a key challenge for a bio-based economy. Because of the complex macromolecular structure, the lignocellulose conversion requires a pretreatment step and an effective hydrolysis. The attention is mainly focusing on the enzymatic hydrolysis, carried out by a tailor-made enzyme cocktail, due to the increasing concerns regarding the environmental impact. A several number of works exploit the autochthonous microbial community involved in lignocellulose decomposition as reservoir of novel lignocellulose-degrading enzymes.

The overall goal of this PhD project was the valorization of selected lignocellulosic biomasses as source of both fermentable sugars and novel biocatalysts for the production of biobased products via fermentation.

Two different lignocellulosic biomasses (the perennial crop *Arundo donax* and the Newspaper Waste – NW – fraction of Municipal Solid Waste) were tested as source of monosaccharides by enzymatic hydrolysis after Ammonia Fiber Expansion (AFEX) or Extractive Ammonia (EA) pretreatment. The ability of the recombinant endocellulase rCelStrep, α-L-arabinofuranosidase rPoAbf and its evolved variant rPoAbf F435Y/Y446F to improve the saccharification yields was evaluated. For the AFEX pretreated *A. donax*, the substitution of rPoAbf F435Y/Y446F to the corresponding enzymatic activity in a mixture of purified enzymes led to obtain higher values (62, 63 and 80% for glucan, xylan and arabinan conversion respectively) than those achieved by using a commercial mixture. The addition of the same enzyme to the commercial mixture enhanced both xylan and arabinan conversions (up to 80%) after 6 days of saccharification. The maximum glucan conversion yield (45.61%) obtained for EA pretreated NW by adding rCelStrep to the commercial mix was higher than or comparable to those reported in the recent manuscripts adopting similar conditions to those used in this study.

The microbial diversity present in three different ecosystems (chipped wood of *Arundo donax, Eucalyptus camaldulensis* and *Populus nigra* subjected to natural biodegradation in open field or underwood) was explored to identify novel lignocellulose-degrading enzymes. By using traditional microorganisms cultivation-based methods, one *Pediococcus acidilactici* strain and five *Streptomyces* strains were identified as producers of novel endo-xylanase(s) and endo-cellulases respectively. The cellulases were tested in the hydrolysis of pretreated *A. donax*. Interestingly, by using the cellulase(s) from the AE-T0-58P (10) strain in substitution of the corresponding activity in a commercial mix, the glucose and xylose yields obtained (4.47±0.5 g/L and 5.87±0.2 g/L respectively after 72 h) were 82% and 85% respectively of the values obtained by using the commercial mix. Moreover, the DNA from the microorganisms adherent to the abovementioned biomasses were subjected to metagenomic analysis. The functional clustering of the open reading frames (ORFs) predicted by the shotgun sequencing of the extracted DNAs showed a prevalence of poorly characterized genes, suggesting the three detected biomasses as potential sources of not yet known proteins. The Glycoside Hydrolases (GHs) abundance in the sample from *P. nigra* (1.85% on total ORFs) was higher than that detected in environments well-known as reservoires of GHs. The in-depth KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway mapping of the putative genes coding for enzymes involved in the polysaccharides hydrolysis confirmed that the three analyzed samples were a valuable source of a full set of (hemi)cellulases and accessory enzymes required for an effective lignocellulose hydrolysis.

Riassunto

Biomasse lignocellulosiche come materia prima per la produzione di biocombustibili e bioprodotti

Il drammatico aumento del prezzo del petrolio, la limitatezza dei combustibili fossili, le crescenti preoccupazioni relative all'impatto ambientale, legate soprattutto alla emissioni di gas serra e al riscaldamento globale, e la sempre crescente attenzione alla salute e alla sicurezza stanno guidando la ricerca verso nuove fonti di energia e processi di produzione che limitino l'uso di sostanze pericolose o nocive. L'utilizzo di materie prime rinnovabili in sostituzione del petrolio potrebbe ridurre al minimo il consumo energetico e la produzione di rifiuti, massimizzando la produttività e la competitività sul mercato (Epicoco et al., 2014).

Ad oggi la produzione di zuccheri fermentabili a scopi industriali utilizza principalmente biomasse commestibili (canna da zucchero, mais, frumento, orzo e barbabietole da zucchero) e il loro sfruttamento ha effetti devastanti sui mercati agricoli, riducendo la quantità di suolo destinata a scopi alimentari e determinandone incrementi di prezzo (Scheidel and Sorman, 2012). Per eliminare questi inconvenienti, sempre maggiore attenzione è rivolta verso le biomasse lignocellulosiche non in competizione con il cibo (colture specifiche su terreni marginali o inquinati, residui agro-industriali o solidi urbani). Tali biomasse lignocellulosiche costituiscono la più abbondante fonte di carbonio disponibile. Lo sviluppo di processi che sfruttino questa risorsa per la produzione di biocombustibili e bioprodotti rappresenta una promettente alternativa all'utilizzo di risorse di origine fossile. Infatti, tali biomasse sono costituite prevalentemente da cellulosa (lunghe catene lineari di unità di glucosio) ed emicellulosa (polisaccaridi ramificati costituiti da differenti esosi e pentosi, tra cui xilosio, arabinosio, mannosio e galattosio), oltre che dalla lignina, che fornisce supporto strutturale, permeabilità e resistenza contro gli attacchi microbici.

L'idrolisi della componente polisaccaridica delle biomasse lignocellulosiche produce monosaccaridi e piccoli oligomeri, materia prima per una vasta gamma di prodotti ad alto valore aggiunto. Il Northwest Laboratory Pacific National (PNNL) e il National Renewable Energy Laboratory (NREL) hanno individuato una serie di composti di interesse industriale prodotti a partire da zuccheri fermentabili. Questi includono l'acido succinico, l'acido gluconico, l'acido aspartico, il glicerolo, il sorbitolo, lo xilitolo (Hermann et al., 2007).

La conversione delle biomasse lignocellulosiche in zuccheri fermentabili è resa difficoltosa dalla loro complessa struttura tridimensionale e richiede, pertanto:

• tecniche di pretrattamento efficaci per la rimozione della lignina aumentando l'accessibilità della componente polisaccaridica;
• metodi di idrolisi eco-friendly ed economicamente competitivi, al fine di ottenere la completa conversione dei polisaccaridi;
• processi di fermentazione da parte di microrganismi in grado di convertire sia esosi che pentosi nei prodotti desiderati.

- **Valorizzazione di colture su terreni marginali. Il caso dell' *Arundo donax***

Sono in corso numerosi studi volti ad individuare colture poliennali su terreni non adatti alle coltivazioni alimentari da utilizzare come fonte di zuccheri fermentabili. In particolare, l' *Arundo donax*, costituita principalmente da cellulosa ed emicellulosa con un basso livello di lignina, appare una delle più interessanti biomasse, data la

sua tolleranza ad una vasta gamma di stress ambientali (elevata salinità del suolo, lunghi periodi di siccità), la capacità di crescere su svariati tipi di terreno, garantendone la crescita anche su suoli degradati o contaminati (Diodato et al., 2009, Diodato et al., 2011; Fagnano et al., 2012).

Nell'ambito di questo progetto di dottorato è stato effettuato il primo studio sull'idrolisi enzimatica della biomassa *Arundo donax* pretrattata mediante tecnologia AFEX (Ammonia Fibre EXpansion). Si tratta di una tecnologia che utilizza sia processi fisici (alta temperatura e pressione) che processi chimici (ammoniaca) per ottenere un efficace pretrattamento. Oltre ad aumentare l'accessibilità della componente polisaccaridica, l' AFEX promuove la decristallizzazione della cellulosa, la parziale depolimerizzazione dell' emicellulosa e riduce la recalcitranza della lignina. Inoltre, la tecnica AFEX permette il recupero quasi totale e la conseguente possibilità di riutilizzo dell'ammoniaca utilizzata durante il processo (Balan et al., 2009).

In primo luogo è stata analizzata la composizione macromolecolare dell' *A. donax* non pretrattato e ne sono state ottimizzate le condizioni di pretrattamento, valutando, per ciascuna condizione, le rese di idrolisi ottenute utilizzando una preparazione enzimatica commerciale. I risultati hanno mostrato che le rese migliori in glucosio e in xilosio sono state ottenute dopo pretrattamento a 130°C per 15 minuti, utilizzando 1 kg di ammoniaca per kg di biomassa secca e con il 60% di umidità. Sulla base di questi dati, tali condizioni sono state utilizzate per i successivi esperimenti di idrolisi. E' stata inoltre valutata la capacità dell'arabinofuranosidasi rPoAbf (prodotta dal fungo *Pleurotus ostreatus* ed espressa in maniera ricombinante in *Pichia pastoris*) (Amore et al., 2012b), del suo mutante evoluto rPoAbf F435Y/Y446F (Giacobbe et al., 2014) e dell' endo-cellulasi rCelStrep (prodotta da *Streptomyces sp. G12* ed espressa in maniera ricombinante in *Escherichia coli*) (Amore et al., 2012a) di migliorare l'idrolisi dell' *A. donax* pretrattato AFEX. L'effetto degli enzimi studiati è stato analizzato utilizzando una miscela di cellulasi purificate (CBHI, CBHII, EGI e βG), endoxilanasi (LX3, LX4) ed enzimi accessori (LarbF e LβX) come miscela di riferimento e sostituendo alternativamente o in combinazione EGI con rCelStrep e LarbF con rPoAbf o rPoAbf F435Y/Y446F. La sostituzione dell'arabinofuranosidasi rPoAbf F435Y/Y446F all'enzima purificato LarbF ha consentito di migliorare la resa di saccarificazione: dopo 72 ore di idrolisi la conversione di glucano, xilano e arabinano è stata di circa 62, 63 e 80%, rispettivamente. Tali valori risultano simili o superiori a quelli (44, 66 e il 55%) ottenuti utilizzando una mix enzimatica commerciale (Ctec3 e Htec3 forniti da Novozymes Cellic®). Inoltre, è stato valutato l'effetto sull'idrolisi dell' *A. donax* pretrattato AFEX dell' aggiunta degli enzimi rPoAbf, rPoAbf F435Y/Y446F e rCelStrep alla mix di enzimi commerciali. Anche in questo caso, è stato dimostrato che l'aggiunta dell'arabinofuranosidasi mutata rPoAbf F435Y/Y446F ha migliorato sia la resa in xilosio che in arabinosio, convertendo l'80% di xilano e arabinano dopo 6 giorni di saccarificazione.

- **Valorizzazione di residui solidi urbani. Il caso della frazione "carta da quotidiani" (Newspaper waste)**

Si stima che entro il 2025 il volume mondiale di rifiuti solidi urbani raggiungerà 2,2 miliardi di tonnellate con un crescente impatto sanitario, sociale ed economico. Per questo motivo sta crescendo l'attenzione nei confronti di modalità di corretta gestione dei rifiuti e di tecnologie volte alla loro valorizzazione. In particolare, grazie ad una disponibilità diffusa e ai bassi costi, le componenti lignocellulosiche da rifiuti

solidi urbani (carta, frazione organica) rappresentano una fonte di zuccheri semplici per la produzione di una vasta gamma di prodotti ad alto valore aggiunto. La carta proveniente da quotidiani ("Newspaper waste") costituisce circa il 14% in peso dei rifiuti solidi urbani prima del riciclo (Subhedar et al. 2015) e rappresenta una delle materie prime rinnovabili più attraenti essendo costituita per circa il 70% da polisaccaridi (Wang et al., 2012). Tuttavia, le differenti tipologie di carta straccia possono contenere non solo lignina, ma anche additivi quali cere, grassi, gomme, polimeri sintetici e resine che impediscono l'idrolisi enzimatica della componente polisaccaridica (Chu and Feng 2013; Kim et al. 2006; Kuhad et al. 2010). Per rimuovere tali composti e migliorare la saccarificazione, tale frazione deve essere sottoposta ad un metodo di pretrattamento appropriato precedente la fase di idrolisi.

E' stata valutata la possibilità di utilizzare il "Newspaper waste" come fonte di zuccheri fermentabili, testando due diversi metodi chimico-fisici di pretrattamento, l'AFEX e l'EA (Extractive Ammonia). Rispetto all'AFEX, la seconda tecnica determina la conversione della celulosa nativa (tipo I) in cellulosa cristallina di tipo III, altamente stabile e più accessibile agli enzimi cellulolitici (da Costa Sousa et al. 2015).

E' stata analizzata la composizione macromolecolare del "Newspaper waste" e ne sono state ottimizzate le condizioni di pretrattamento AFEX, valutando, per ciascuna condizione, le rese di idrolisi ottenute utilizzando una preparazione enzimatica commerciale. In base ai dati ottenuti e al consumo energetico connesso alle differenti condizioni prese in esame, per i successivi esperimenti di idrolisi sono stati scelti i seguenti parametri: 2.8 kg ammoniaca per kg biomassa secca, 65°C e 10.7% di umidità con un tempo di reazione di 15 minuti. Per il trattamento EA sono state scelte le seguenti condizioni: temperatura di reazione 120°C; 3 kg di ammoniaca per Kg biomassa secca; contenuto di umidità 10%; tempo di reazione 15 minuti.

Inoltre, sulla base dei promettenti risultati ottenuti per l'idrolisi dell' *A. donax* pretrattato AFEX (Giacobbe et al., 2015), gli enzimi ricombinanti rCelStrep, rPoAbf e la sua variante migliorata rPoAbf F435Y/Y446F sono stati utilizzati nella saccarificazione del "Newspaper Waste" pretrattato AFEX ed EA. In particolare, la stessa miscela di enzimi purificati utilizzata per gli esperimenti su *A. donax*, contenente cellulasi, xilanasi ed attività enzimatiche accessorie, è stata scelta come mix di riferimento e rCelStrep, rPoAbf o la sua variante sono stati sostituiti rispettivamente a EGI e Larb. I risultati hanno mostrato che nessuna di queste mix enzimatiche è efficace nell'idrolisi del "Newspaper Waste" dopo pretrattamento AFEX o EA. D'altra parte, quando gli enzimi rCelStrep, rPoAbf e rPoAbf F435Y/Y446F sono stati aggiunti alla mix di enzimi commericali (Ctec3 e Htec3 forniti da Novozymes Cellic®), è stato dimostrato che aggiungendo rPoAbf la resa totale di conversione della componente polisaccaridica raggiungeva il 37.32% per il "Newspaper Waste" pretrattato AFEX, mentre aggiungendo rCelStrep la massima resa di conversione ottenuta era del 40.80% per il "Newspaper Waste" pretrattato EA. In quest'ultimo caso, si otteneva anche la resa massima di conversione del glucano (45.61%). Tale valore è superiore o comparabile a quelli riportati nei più recenti lavori che adottano condizioni di idrolisi simili a quelle usate in questo studio (Wang et al., 2012; Wu et al., 2014; Subhedanar et al., 2014).

Biomasse lignocellulosiche come fonte di nuovi biocatalizzatori

Diversi habitat (suolo, compost, materiale vegetale in decomposizione, cumuli di rifiuti forestali, fanghi di depurazione, sorgenti di acqua calda, rumine, feci di

animali) sono stati studiati come fonti di microrganismi aerobi ed anaerobi produttori di differenti enzimi di interesse industriale. In particolare, le biomasse lignocellulosiche rappresentano un ecosistema complesso in cui le condizioni ambientali possono influenzare la biodiversità, favorendo le comunità microbiche autoctone sugli altri microrganismi grazie alla loro capacità di produrre enzimi coinvolti nella degradazione della (emi)cellulosa.

In collaborazione con il team della prof.ssa Olimpia Pepe del Dipartimento di Agricoltura dell'Università degli Studi di Napoli " Federico II ", è stata analizzata la diversità microbica presente negli ecosistemi naturali costituiti dalle biomasse *Arundo donax, Eucalyptus camaldulensis* e *Populus nigra*, utilizzando sia un approccio tradizionale di isolamento e coltura diretta dei microrganismi, sia un approccio metagenomico con l'obiettivo di identificare nuovi biocatalizzatori coinvolti nella degradazione della lignocellulosa. In particolare, del cippato di *A. donax, E. camaldulensis* e *P. nigra* è stato sottoposto a biodegradazione naturale in campo aperto e in sottobosco per 180 giorni, al fine di aumentare la biodiversità microbica autoctona. I campioni sono stati raccolti subito dopo la preparazione (T0) e dopo 45, 90, 135 e 180 giorni di biodegradazione (T1, T2, T3 e T4 rispettivamente).

- **Identificazione di nuove (emi)cellulasi mediante tecniche tradizionali di coltura diretta dei microrganismi**

A partire dalle succitate biomasse sono state isolate 1157 colonie (33% eucarioti e 67% procarioti). Tutti i microrganismi isolati sono stati testati per le diverse attività enzimatiche target (endo- ed eso-cellulasica, emicellulasiche ed attività accessorie) su mezzo solido contenente substrati specifici (carbossimetilcellulosa - CMC - , xilano, pectina e Avicel). I 40 batteri che mostravano il più esteso alone di attività su CMC agar e multi-attività enzimatiche sono stati identificati mediante sequenziamento dei geni codificanti il 16S rRNA. Mediante screening quantitativi in mezzo liquido, il ceppo ET2C-75A - *Pediococcus acidilactici* si è rivelato il maggior produttore di attività endo-1,4-beta-glucanasica (0,32 U/ml dopo 12 ore di incubazione) e xilanasica (1 U/mL dal 3 al 9 ° giorno di crescita in terreno liquido contenente xilano come unica fonte di carbonio). Attraverso analisi di zimografia accoppiate ad analisi di proteomica (in collaborazione con il gruppo della Prof.ssa Birolo del Dipartimento di Scienze Chimiche dell' Università di Napoli "Federico II") sono stati identificati peptidi appartenenti a tre diverse endo-xilanasi prodotte da *Cellulomonas flavigena*, *Actinosynnema mirum* e *Actinoplanes sp.* Questi risultati hanno suggerito che le xilanasi prodotte dal microrganismo selezionato *P. acidilactici* non sono ancora state depositate in banca dati.

E' stato effettuato un ulteriore screening sulle già menzionate 1157 colonie isolate e purificate al fine di individuare ceppi di Attinomiceti con potenzialità di degradazione della biomassa lignocellulosica. Infatti, come ampiamente riportato in letteratura, molti Attinomiceti possono essere utilizzati nella bioconversione dei rifiuti agricoli e urbani grazie alla loro capacità di sintetizzare biocatalizzatori necessari alla completa idrolisi delle componenti recalcitranti delle biomasse lignocellulosiche. I ventiquattro ceppi selezionati, identificati mediante sequenziamento dei geni codificanti il 16S rRNA, appartenevano al genere *Streptomyces*. Uno screening quantitativo in mezzo liquido contenete CMC come unica fonte di carbonio ha permesso di selezionare i migliori 5 ceppi produttori di attività endo-cellulasica. E' stata effettuata una parziale caratterizzazione degli enzimi di interesse, valutandone attività residua a tre diverse temperature di conservazione, optimum di pH, optimum di temperatura e termoresistenza. I risultati hanno mostrato che nessuna condizione

di conservazione selezionata era adatta all'endo-cellulasi prodotta dal ceppo AE-T2C-612P (9), mentre gli enzimi prodotti dagli altri ceppi mostravano un'elevata stabilità fino a 30 giorni di conservazione a -20° e a -80°C. In buffer sodio citrato, le attività cellulasiche prodotte da tutti i ceppi di *Streptomyces* analizzati mostravano un optimum di pH pari a 5.0 e un optimum di temperatura pari a 50°C. Queste condizioni sono state utilizzate per testare le suddette attività enzimatiche nella bioconversione di *A. donax* pretrattato. Gli esperimenti di saccarificazione sono stati effettuati utilizzando un cocktail enzimatico commerciale come riferimento e sostituendo alternativamente i surnatanti contenenti le endo-cellulasi prodotte dai ceppi di *Streptomyces* alla corrispondente attività enzimatica nella mix commerciale. I risultati hanno mostrato che dopo 72 di idrolisi usando la mix contenente il surnatante proveniente da una coltura del ceppo AE-T0-58P (10) si otteneva una resa in glucosio e xilosio di 4,47 ± 0,5 g/L e 5,87 ± 0,2 g/L rispettivamente. È interessante notare che questi valori rappresentano rispettivamente l'82% e 85% delle corrispondenti rese ottenute utilizzando la miscela commerciale di riferimento.

- **Identificazione di nuove (emi)cellulasi mediante approccio metagenomico**

Come ampiamente dimostrato, solo una piccolissima percentuale di microrganismi presenti in natura (non superiore all'1%) può essere coltivata in condizioni standard di laboratorio (Whitman et al., 1998). Per superare questo limite, l'estrazione e l'analisi del materiale genetico direttamente da campioni ambientali (definito "metagenomica") rappresenta una valida alternativa. (Sharma et al., 2008). Possono essere utilizzate due differenti tipologie di screening su "DNA ambientale": un approccio basato sull'analisi funzionale di librerie di espressione o un approccio legato al sequenziamento dell'intero materiale genetico estratto. Il DNA estratto dai microrganismi sviluppati sulle biomasse lignocellulosiche summenzionate lasciate a biodeteriorare naturalmente è stato sia utilizzato per la costruzione di una library fosmidica di espressione che sottoposto a sequenziamento, al fine di individuare nuovi biocatalizzatori convolti nella degradazione della lignocellulosa. Sebbene l'approccio funzionale non abbia prodotto i risultati sperati, i dati relativi al sequenziamento del DNA della biomassa microbica sviluppata nei cumuli di cippato di *Arundo donax, Eucalyptus camaldulensis* e *Populus nigra* dopo 135 giorni di biodegradazione in sottobosco (T3ADSB, T3ESB and T3PSB rispettivamente) sono apparsi molto interessanti. In particolare, la prevalenza di putative ORF classificate nel database egg-NOG (evolutionary genealogy of genes: Non-supervised Orthologous Groups) come scarsamente caratterizzate suggerisce che tali biomasse possano essere potenziale fonte di geni non ancora noti. Inoltre, la percentuale di geni legati al pathway metabolico dei carboidrati, individuati mediante confronto con il database KEGG (Kyoto Encyclopedia of Genes and Genomes), è superiore o simile a quella rilevata in metagenomi provenienti da campioni con una ben nota capacità di degradazione della lignocellulosa. Questo risultato conferma l'elevata potenzialità dei campioni analizzati di esprimere geni coinvolti nella biotrasformazione delle biomasse lignocellulosiche. Inoltre, le ORF individuate nei tre metagenomi sono state confrontate col database CAZy (Carbohydrate Active enZymes) che raccoglie tutti gli enzimi attivi sui carboidrati. E' interessante notare la presenza di ORF codificanti putative monoossigenasi coinvolte nella lisi dei polisaccaridi (LPMO), cui si rivolge sempre maggior attenzione, a causa della loro capacità di depolimerizzare la componente recalcitrante insolubile della cellulosa, incrementando l'efficienza di saccarificazione della lignocellulosa (Beeson et al., 2015). La maggior parte dei

putativi CAZy rilevati nei tre campioni erano identificate come glicosil-idrolasi (GH), con percentuali rispetto al numero totale delle ORF (0.68%, 0.34% e 1.85% in T3ADSB, T3ESB e T3PSB rispettivamente) superiori a quelle rilevate in metagenomi da suolo, fanghi e ambienti marini o lacustri. In più, la percentuale di GH rilevata nel campione T3PSB era superiore anche a quelle riportate per il microbiota di organismi quali elefanti, lumache, mucce e termiti (Ilmberger et al, 2014; Cardoso et al, 2012; Brulc et al, 2009; Warnecke et al, 2007), ben noti per la loro potenzialità nell'esprimere un elevato numero di enzimi coinvolti nella degradazione dell' (emi)cellulosa. La maggior parte delle GH per ciascun campione (~ 26% per T3ADSB, ~ 22% per T3ESB e ~ 24% per T3PSB) erano coinvolti nell'idrolisi degli oligosaccaridi, con una predominanza di cellulasi per il campione T3ADSB (6.2%), di emicellulasi per il campione T3PSB (5.0%) e di enzimi coinvolti nell'idrolisi delle catene laterali di polisaccaridi per il campione T3ESB (4.3%). Tali dati hanno confermato che i metagenomi provenienti da ciascuna delle tre biomasse lasciate naturalmente a biodeteriorare sono una fonte preziosa di tutti gli enzimi (cellulasi, emicellulasi ed enzimi accessori) necessari ad una efficace saccarificazione della lignocellulosa.

List of abbreviations

AA Auxiliary Activities enzyme
AFEX Ammonia Fiber EXpansion
bp base pair
BSA Bovin serum albumin
CAZy / CAZyme Carbohydrate-Active Enzymes
CBHI cellobiohydrolase I
CBHII cellobiohydrolase II
CBM Carbohydrate-binding modules
CBP Consolidated BioProcessing
CE Carbohydrate Esterases
CMC CarboxyMethylCellulose
DNS 3,5-dinitrosalicylic acid
EA Extractive Ammonia
EC Enzyme Commission number
eDNA environmental DNA
eggNOG evolutionary genealogy of genes: Non-supervised Orthologous Groups
GH Glycoside Hydrolases
GT Glycosiltransferase
HPLC High Performance Liquid Chromatography
KEGG Kyoto Encyclopedia of Genes and Genomes
LC-MS / MS liquid chromatography-mass spectrometry/mass spectrometry
LPMO Lytic Polysaccharide MonoOxygenases
MSW Municipal Solid Waste
NCBInr National Center for Biotechnology Information – non redundant database
NREL National Renewable Energy Laboratory
NT NCBI National Center for Biotechnology Information – nucleotide database
NW Newspaper Waste
OD Optical density
ORF Open Reading Frame
PCA Plate Count Agar
PCR Polymerase Chain Reaction
PL Polysaccharide Lyases
PVDF Polyvinylidene fluoride
Q-TOF Quadrupole-time-of-flight
SD Synthetic Defined medium
SDS-PAGE Sodium Dodecyl Sulphate - PolyAcrylamide Gel Electrophoresis
SHF Separate Hydrolysis and Fermentation
SSCF Simultaneous Saccharification and Co-Fermentation
SSF Simultaneous Saccharification and Fermentation
T3ADSB *Arundo donax* after 135 days of natural biodeterioration in underwood
T3ESB *Eucalyptus camaldulensis* after 135 days of natural biodeterioration in underwood
T3PSB *Populus nigra* after 135 days of natural biodeterioration in underwood
TFA Trifluoroacetic acid
YP bacto-yeast extract and bacto-peptone medium

Chapter 1

Introduction

1.1 Lignocellulose as raw material for biochemicals production

The dramatic increase in the price of petroleum, the finite nature of fossil fuels, the increasing concerns regarding environmental impact, especially related to the greenhouse gas emissions and to the global warming, and the health and safety considerations are forcing the search for new energy sources and production processes with the aims of reducing or eliminating the use of dangerous or hazardous substances. The use of renewable raw materials in substitution of petroleum minimizes energy consumption and waste generation maximizing the productivity and competitiveness in the marketplace. Nowadays the edible biomasses (such as sugarcane, corn, wheat, barley and sugarbeets) are still largely used as feedstock for fermentable sugars for other purposes than food or feed having devastating effects on agricultural markets, reducing food production and consequently increasing food prices (Scheidel and Sorman, 2012). To overcome this drawbacks, increasing attention is focused on the use of raw materials non in competition with food.

Lignocellulosic biomasses constitute the Earth's most abundant repository of carbon, thus the future development of processes that exploit this resource for industrial applications represents a key challenge for a bio-based economy. Lignocellulose consists of cellulose, hemicellulose, lignin, organic extractives (mixture of different organic compounds) and some inorganic components (figure 1).

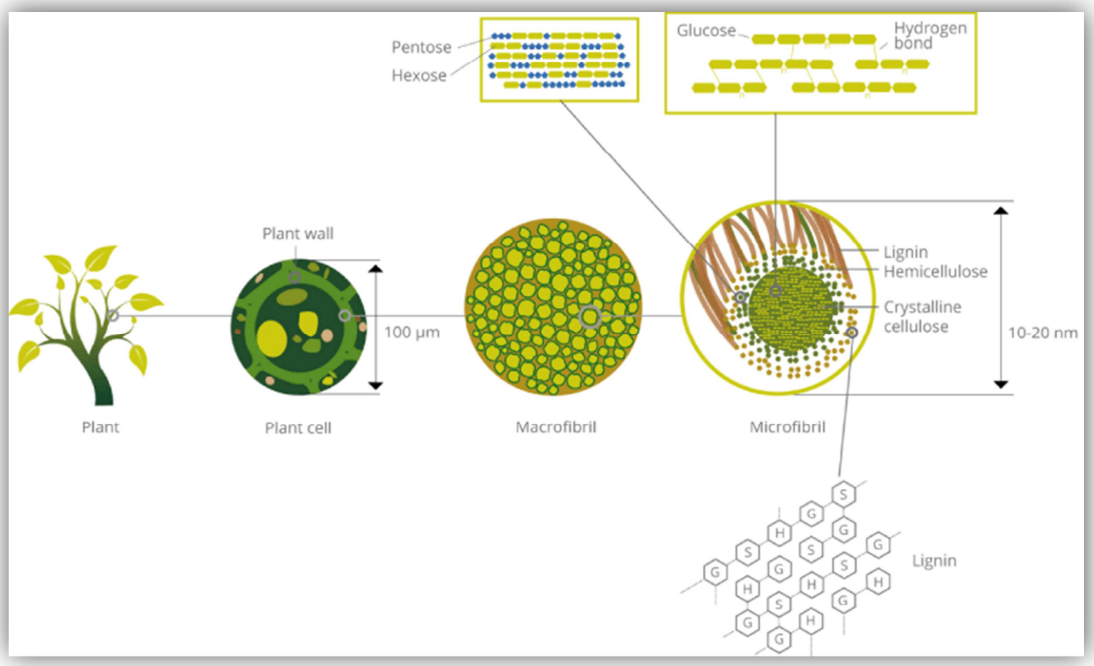


Figure 1. Schmatic structure of lignocellulose. The hexagons denote the lignin subunits *p-coumaryl alcohol (H), coniferyl alcohol (G) and sinapyl alcohol (S). From "Streffer, F. Lignocellulose to Biogas and other Products. JSM Biotechnol Bioeng 2(1): 1023.2014." ©Creative Commons Attribution 4.0 International License*

Cellulose, hemicellulose and lignin constitute more than 75% of the vegetal material and are composed of organic polymers of high molecular weight. In particular, cellulose is a crystalline and amorphous organic compound, consisting of a linear chain of several hundred to over ten thousand β(1→4) linked units of glucose. Hemicelluloses are branched polysaccharides, accounting for 25-35% of dry wood,

with a random and amorphous structure with little strength, consisting in a mixture of about 500 - 3000 units of various monosaccharides, such as xylose and arabinose (pentose), mannose and galactose (hexose) and glucoronic acid. Lignin is present in the cell walls, giving structural support, impermeability and resistance against microbial attack and oxidative stress. It has a highly complex structure made up of three different phenyl propane monomers joined together by non hydrolyzable linkages.

A wide variety of different lignocellulose biomasses have been tested and used to produce fermentable sugars.

Several efforts are under way to use lands not suitable for traditional food crops for growing dedicated energy crops, such as giant reed (*Arundo donax*), mischantus and switchgrass. In Mediterranean environments, hilly areas are considered not appropriate for the traditional cereal production because yield and gross income are very low and because the traditional cropping system (deep soil tillage at the end of August and sowing at November) causes extreme vulnerability to soil erosion (Diodato et al., 2009, Diodato et al., 2011; Fagnano et al., 2012). In these conditions, perennial biomass crops are proved to reduce soil erosion and to increase potential gross income of farmers (Fagnano et al., 2015) with favourable environmental impacts (Forte et al., 2015). Moreover, these crops are also suggested for other areas not suitable for food crops such as polluted soils since it acts as a phyto-remediating agent (Fiorentino et al., 2010; Fiorentino et al., 2013).

Moreover, lignocellulose agricultural wastes (such as sugarcane baggase, rice hull, timber species, willow, salix, switch grass, softwood, rice straw, wheat straw) or industrial wastes (such as Brewers' spent grain) could be utilized for the production of a number of value added products, due to their chemical composition based on polysaccharides and other compounds of interest. These wastes are produced every year in large quantities and their incineration or disposal to the soil or landfill causes serious environmental problems. Furthermore, the price of these materials is significantly lower than the price of vegetable oil, corn and sugarcane.

In addition to these biomasses, the lignocellulosic fractions of Municipal Solid Waste (MSW), such as paper or organic fractions, can be used as renewable resources of a wide range of high added value products (Liguori, Amore and Faraco, 2013) due to the widespread availability, low pricing and suitability for most conversion technologies. In fact, the world MSW volume is expected to reach 2.2 billion tonnes by 2025 and improperly managed urban solid waste is one of the main causes of environmental pollution and a serious health hazard, due to contamination of groundwater and surface water by leachate, as well as air pollution by their burning.

The amounts of carbohydrates and lignin are highly variable from one biomass to another and depends on different conditions (such as age of plant, stage of growth) (Mussatto and Teixeira, 2010). Average values of the main components in some non-edible lignocellulose biomasses are shown in Table 1 (nee'Nigam, Gupta and Anthwal, 2009; Liguori et al., 2015; Ioelovich, 2015).

Table 1. Main components of different lignocellulosic biomasses.

Lignocellulose waste	Cellulose (wt %)	Hemicellulose (wt %)	Lignin (wt %)
Barley straw	33.8	21.9	13.8
Corn cobs	33.7	31.9	6.1
Corn stalks	35.0	16.8	7.0
Cotton stalks	58.5	14.4	21.5

Oat straw	39.4	27.1	17.5
Rice straw	36.2	19.0	9.9
Rye straw	37.6	30.5	19.0
Soya stalks	34.5	24.8	19.8
Sugarcane bagasse	40.0	27.0	10.0
Sunflower stalks	42.1	29.7	13.4
Wheat straw	32.9	24.0	8.9
Brewers' Spent Grain	14.4	34.2	3.9
Office paper	62	5	1
Newspaper waste	38	15	21

The catalytic hydrolysis of lignocellulose produces monosaccharides and small oligomers, raw material for a wide range of high value-added products. A lot of different chemical and biological processes (i.e. gasification, liquefaction, pyrolysis, hydrogenation and fermentation) and/or cascade reactions were performed to obtain potentially industrial relevant chemicals from sugars. The Pacific Northwest National Laboratory (PNNL) and National Renewable Energy Laboratory (NREL) identified 12 building blocks produced through sugars conversion which could be used to obtain a variety of high value bio-based products. These include sorbitol and fructose (used in food and non-food industry), alkyl polyglucosides (applied as surfactant), polylactic acid (mainly used in food packaging), succinic acid (precursor to polymers, resins, and solvents), 5-hydroxymethyl-2-furfural (a versatile molecule involved in the production of solvents, transportation fuels and polymers) (Lichtenthaler and Peters, 2004; Hermann and Patel, 2007). Nowadays the bioethanol is the main bulk chemical produced from lignocellulose biomasses.

1.2 Main steps of the production processes

Lignocellulose is difficult to be converted into fermentable sugars, due to the complex three-dimensional structure of interwoven polymers forming a highly resistant matrix protecting plants from biotic and abiotic aggressions. Therefore, in order to develop a competitive process for lignocellulose conversion, the following steps are required:

- effective and inexpensive pretreatment methods to disarray the recalcitrant lignin increasing the polysaccharides accessibility for the further saccharification;
- economically competitive enzymatic hydrolysis, more eco-friendly than thermal or chemical saccharification, in order to achieve the conversion of whole polysaccharide component into fermentable sugars;
- fermentation of sugars from previous step into target products by microorganisms capable of converting both hexose and pentose monosaccharides.

1.2.1 Pretreatment methods

The goals of the pretreatment step include: disrupting of recalcitrant lignocellulose structures making (hemi)cellulose more accessible for further hydrolysis; avoiding of the degradation of sugars mainly those derived from hemicellulose; minimizing of the inhibitors formation for subsequent fermentation step. A lot of different pretreatment technologies have been developed and the "best"

strategy to be adopted depends on many factors such as type of lignocellulosic biomass and desired products, as well as the economic impact on overall bioconversion process (Eggeman and Elander, 2005). The pretreatment techniques can be classified into different categories according to various criteria (Xu and Huang, 2014). It is common to differentiate the methods based on the principal mechanism acting during pretreatment. As such, the methods can be classified as mechanical, chemical, physicochemical, biological or various combinations thereof.

Mechanical pretreatment is generally performed before a following processing step and encompass chipping, milling or grinding (to reduce the particle size) and ultrasonic pretreatment (to increase the polysaccharides accessibility) probably due to the breaking of hydrogen bonds of the crystalline cellulose.

The pretreatments achieved by **chemical** reactions or **physicochemical** mechanisms for disruption of the biomass structure are briefly described below.
The Steam Explosion is a physicochemical method carried out with high-pressure saturated steam followed by a swiftly reduction of pression causing an explosive decompression. The addition of SO_2 or H_2SO_4 significantly increases the hemicellulose sugar yields.
The hydrothermal pretreatment Liquid Hot Water is carried out by using high temperature and pressure and results in hemicelluloses solubilization and cellulose digestibility increasing. However, this method requires remarkable amount of energy because of high pressures and large amounts of water supplied to the system.
The acid pretreatments encompass Dilute Acid Pretreatment and Organosolv. Dilute Acid Pretreatment is usually performed over a temperature range of 120 to 210 °C, with acid concentration typically less than 4% wt and residence time from several minutes to an hour reducing significantly the lignocellulosic recalcitrance (Taherzadeh and Karimi, 2007). In the Organosolv pretreatment, numerous organic or aqueous solvent mixtures, such as ethanol and methanol, are used at high cooking temperature up to 200°C for 30–90 min with a solvent concentration of 35 – 70 % (w/v) and an acid pH (2-4). The organic solvents are expensive, might inhibit the enzymatic hydrolysis and further fermentation and causes high energy consumption (Zhao, Cheng and Liu, 2009).
The alkaline methods can be divided into two groups based on the chemical used, sodium or calcium hydroxide and ammonia. The calcium and sodium hydroxide pretreatments are carried out at low temperatures and pressures, even at ambient conditions, but the reaction times can be long up to days. However, the hydroxide can be converted into irrecoverable salts and the incorporation of the salts into the biomass is a relevant issue (Kim, Lee and Kim, 2016). The Ammonia Fiber EXpansion (AFEX) pretreatment is a dry to dry physicochemical process that utilizes high temperatures and pressures combined with chemical (ammonia) processes to achieve minimal carbohydrate degradation and negligible modifications in total polysaccharide composition (Balan et al., 2009).
The oxidative delignification is carried out by the biomasses treatment with an oxidizing agent (such as hydrogen peroxide, ozone, oxygen or air) that reacts with the aromatic ring. However, a substantial part of the hemicellulose is degraded and can no longer be used for sugar production.
CO_2 explosion and mechanical/alkaline pretreatment are other tecniques that combine mechanical and chemical action in order to enhance the polysaccharides component accessibility.

The **biological** delignification approach includes the use of fungi (such as white-rot fungi, brown-rot fungi and soft-rot fungi) and bacteria able to attack and degrade lignin, thus making easier access to (hemi)cellulose. These methods are very cost effective and environmentally friendly, do not require high energy input and produce low toxic materials such as furfural and hydroxymethylfurfural. However, the rate of biological hydrolysis is usually very low, so this pretreatment requires long residence times (Narayanaswamy et al., 2013).

1.2.2 Hydrolysis: chemical versus enzymatic saccharification

Currently, there are two major ways to convert (hemi)cellulose into fermentable sugars: chemical or enzymatic saccharification. The chemical catalytic hydrolysis, generally carried out by using acid compounds (such as H_2SO_4), is a fast acting reaction that requires low residence time. However, the residual acid at the end of process is only neutralized and usually not recycled causing economic and environmental issues. The enzymatic hydrolysis is carried out by means of different enzymes that act in synergism to break the glycosidic bonds in (hemi)cellulose. Enzymatic hydrolysis do not require corrosion resistant processing equipment and produce low acid waste with high yields. This technique is higher in selectivity and requires lower energy costs and milder operating conditions than chemical processes (Binod et al., 2011). Because of the combination of higher conversion, cheaper utility costs and less environmental impact, enzymatic hydrolysis is a very promising lignocellulose saccharification method.

However, the main limitations to the industrial production of bio-based chemicals via enzymatic catalysis from lignocellulose is related to the high cost of the enzymes currently employed in biomass conversion, accounting for more than 20% of the cost of the overall process (Stephen, Mabee and Saddler, 2012), their instability and low activity under the required operating conditions.

In order to develop effective and cost competitive enzymatic hydrolysis processes, different strategies are pursued such as:

1. the identification of novel biocatalysts from new microorganisms by a traditional cultivation based approach or by metagenomic tecniques;
2. the rational design and/or directed evolution of available enzymes in order to obtain improved variants;
3. the immobilization of industrial enzymes in order to enhance stability and the reusability of enzymes;
4. the development of improved processes for enzymes production;
5. the on-site production of target enzymes.

- **Enzymes involved in lignocelluloses conversion**

Enzymatic hydrolysis is influenced by both structural features of (hemi)celluloses and mode of enzyme action. Due to the complexity of the structure, the mechanism of hydrolysis of these substrates is still not fully understood, although detailed knowledge of some aspects of enzyme structure, enzyme molecular properties and the ultrastructure of (hemi)celluloses have been obtained through extensive study over the last few decades.

As reported in several previous studies, a tailor made enzyme cocktail is required to achieve an efficient saccharification (Gao et al., 2010a; Gao et al., 2011).

The cellulases are enzymes which hydrolyze the β (1→4) glycosidic bonds. The endoglucanases (EC 3.2.1.4) carry out the degradation of amorphous cellulose, by

randomly hydrolyzing of the internal glycosidic bonds. The cellobiohydrolases are exocellulases which cut two to four units from reducing ends - CBHI (EC 3.2.1.-) - or non-reducing ends - CBHII (EC 3.2.1.91) - of cellulose. The cooperation between cellobiohydrolases and endoglucanases is crucial for an efficient hydrolysis of the cellulose. The β-glucosidases (EC 3.2.1.21) carry out the degradation of cellobiose and other cellodextrins playing a key role in enzymatic lignocellulose degradation because of the high cellobiose level produced by endoglucanases and cellobiohydrolases. Recently, there is a growing attention to auxiliary enzymes acting towards cellulose by a non-hydrolytic mechanism of depolymerization. The lytic polysaccharides monooxygenases (LPMOs) are the most promising class due to their capability of enhancing the efficiency of lignocellulose biomass degradation by acting on recalcitrant to cellulases polysaccharides from highly crystalline cellulose Beeson et al., 2015; Dimarogona et al., 2013). The LPMOs are involved in the depolymerization of the highly crystalline cellulose carrying out the cleavages of the internal glycosidic bond through hydroxylation at C1, C4 or C6 carbon and subsequent release of oxidized soluble oligosaccharides. This mechanism provides new sites for the action of cellobiohydrolases and β-glucosidases, enhancing the efficiency of cellulose depolymerization. Critical for the region-selectivi hydroxylation of crystalline are surface expsosed catalytic copper atoms.

The hemicelluloses are polysaccharides made by different glyco-units and glycosidic bonds towards which several enzymes are specific. An extensive (but non-exhaustive) list of enzymes involved in hemicelluloses degradation is reported below. The endo-xylanases (EC 3.2.1.8) carry out the degradation of (glucurono)(arabino)xylan, cleaving the backbone glucosidic bonds in xylan. The β-xylosidases (EC 3.2.1.37) act on the xylooligosaccharides, after the action of endo-xylanases towards xylan. β-glucanases (EC 3.2.1.-) hydrolyze β(1→3), (1→4), or (1→6) glucan. Mannanase (EC 3.2.1.78) is involved in degradation of galacto-gluco mannans. Xyloglucan hydrolases (EC 3.2.1.150, 151, 155) belong to the xyloglucan transferase/hydrolase and carry out the hydrolysis of xyloglucan. The degradation of pectic is carried out by several pectinolytic enzymes, such as polygalacturonases (EC 3.2.1.15, 67, 82), pectin lyases, pectate lyases (EC 4.2.2.2, 6, 9, 10) and pectin methyl esterases (EC 3.1.1.11). Acetyl and feruloyl ester substituents in hemicelluloses are removed by acetyl xylan esterase (EC 3.1.1.72) and feruloyl esterase (EC 3.1.1.73), respectively. Alpha-L-arabinofuranosidases (EC 3.2.1.55), alpha-galactosidases (EC 3.2.1.22), alpha-glucuronidases (EC 3.2.1.139) carry out the removal of arabinose, galactose and glucuronoyl substituents in hemicelluloses.

- **Lignocellulosic biomasses as source of novel biocatalysts**

The lignocellulose-degrading enzymes are effectively produced by different microorganisms, such as bacteria belonging to *Clostridium, Cellulomonas, Bacillus, Thermomonospora, Ruminococcus, Erwinia, Acetovibrio, Microbispora,* and *Streptomyces* (Ventorino et al., 2015) and fungi (*Sclerotium rolfsii, P. chrysosporium, Trichoderma spp., Aspergillus spp., Schizophyllum spp., Talaromyces spp.* and *Penicilium spp.*) (Dos Reis et al., 2003; Jorgensen et al., 2003; Balat, 2011). Nevertheless the discovery of novel biomass-degrading biocatalysts to use in bioconversion processes could decrease costs and increase economic sustainability in the industrial applications. Recently, several studies showed the relevance of the research for novel lignocellulolytic enzymes in samples from environments which favor the growth of microorganisms able to degrade lignocellulose, such as microbiota from guts of plant-fed elephant (Ilmberger et al., 2014), invasive snail

(Cardoso et al., 2012), cow (Brulc et al., 2009) and termite (Warnecke et al., 2007). It is also worth of note that lignocellulosic materials are efficiently decomposed by cooperation of autochthonous microbial communities and the environmental conditions, such as pH, salinity, temperature, pressure and radiation, can have a selective pressure on the biodiversity of living microorganisms (Feng et al., 2011). As a consequence, autochthonous microbial communities may prevail over other microorganisms because they possess enzymes able to degrade complex molecules such as cellulose and hemicellulose. Therefore, an increasing number of works exploit these ecosystems, such as compost obtained from agro-industrial wastes (Amore et al. 2012a, Dougherty et al., 2012) and chipped wood after natural biodegradation (Ventorino et al., 2015), showing that this feedstock can be considered a reservoir of novel microorganisms and genes encoding lignocellulose-degrading enzymes to upgrade the feasibility of lignocellulose conversion.

1.2.3 Fermentation strategies

The final step of fuel and biochemicals production processes from lignocellulose is the fermentation of sugars obtained from the previous hydrolysis step. The main choice that can be followed for this step are shown in figure 2 and described below.

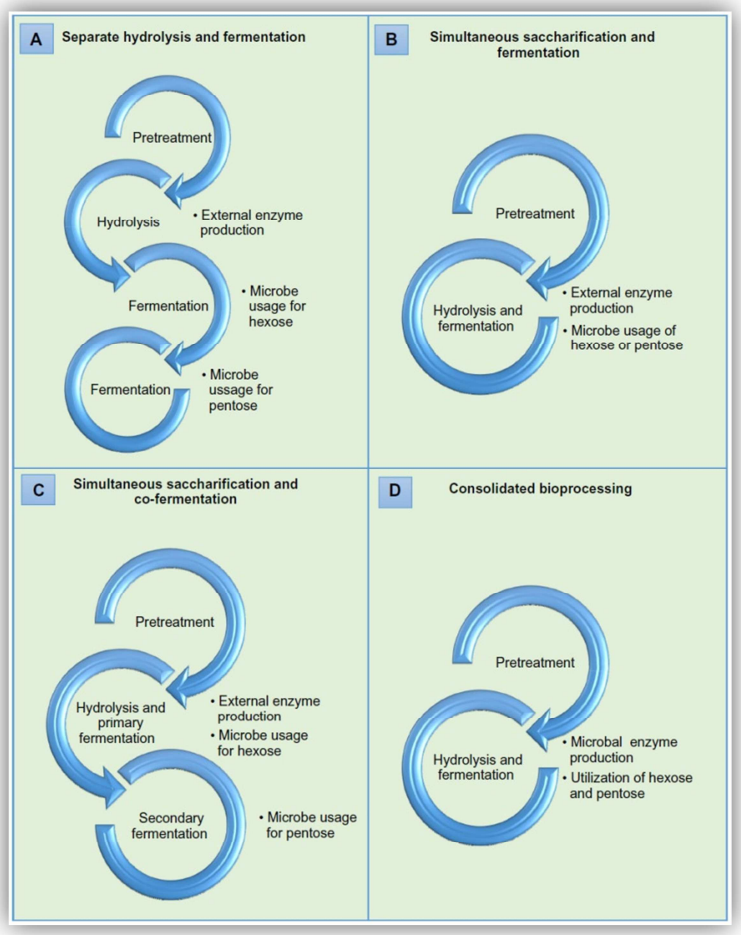


Figure 2. Bioprocessing approaches. (A) separate hydrolysis and fermentation; (B) simultaneous saccharification and fermentation; (C) simultaneous saccharification and co-fermentation; and (D) consolidated bioprocessing. *From "Mbaneme-Smith V and Chinn MS. Consolidated bioprocessing for biofuel production: recent advances. Energy and Emission Control Technologies. 2014." ©Creative Commons Attribution - Non Commercial (unported, v3.0) License*

In Separate Hydrolysis and Fermentation (SHF) the saccharification and the fermentation take place in two separate vessels at their respective optimal conditions, such as pH or temperature. In Simultaneous Saccharification and Fermentation (SSF) a unique reactor is used for both enzymatic hydrolysis and fermentation, reducing the capital cost, increasing the hydrolysis rate and having no losses in fermentable monosaccharides liberated during the hydrolysis (Olofsson, Bertilsson and Lidén, 2008). The effective bioconversion of lignocellulosic biomasses containing a high percentage of hemicelluloses released after pretreatment can be achieved by SSCF, namely Simultaneous Saccharification and Co-Fermentation of pentose and hexoses (Olofsson, Bertilsson and Lidén, 2008). Nowadays, several studies are focusing on the development of Consolidated BioProcessing (CBP) systems in which a single engineered microorganism is able to directly convert (hemi)cellulose into biofuels and value-added products using its own enzymatic machinery, avoiding the requirement of external enzymes addition.

1.3 The Biorefinery Concept

According to the definition performed by the IEA Bioenergy Task 42, “Biorefining is the sustainable processing of biomass into a spectrum of marketable products and energy”. The biorefinery concept is analogous to petroleum-based refinery (which produces fuels and products from petroleum) and is related to a facility, a process, a plant, or even a cluster of facilities to convert biomass resources into building blocks (carbohydrates, proteins, triglycerides) for the production of biofuels and high-value added products. Nowadays, most of the biofuels and biochemicals are produced in single chains and not within a biorefinery concept, usually requiring materials in competition with food and feed. Contrariwise, a biorefinery is assessed for the entire value chain on their environmental, economic, and social sustainability covering the whole life cycle (construction— operation— dismantling) and can use all kinds of biomass from forestry, agriculture, aquaculture, and residues from industry and households including wood, agricultural crops, organic residues (both plant and animal derived), forest residues, and aquatic biomass (algae and seaweeds). The figure 3 shows a schematic overview of the biorefinery concept. Biorefineries are expected to contribute to an increased competitiveness and wealth of the countries by responding to the need for supplying a wide range of biobased products and energy in an economically, socially, and environmentally sustainable manner.

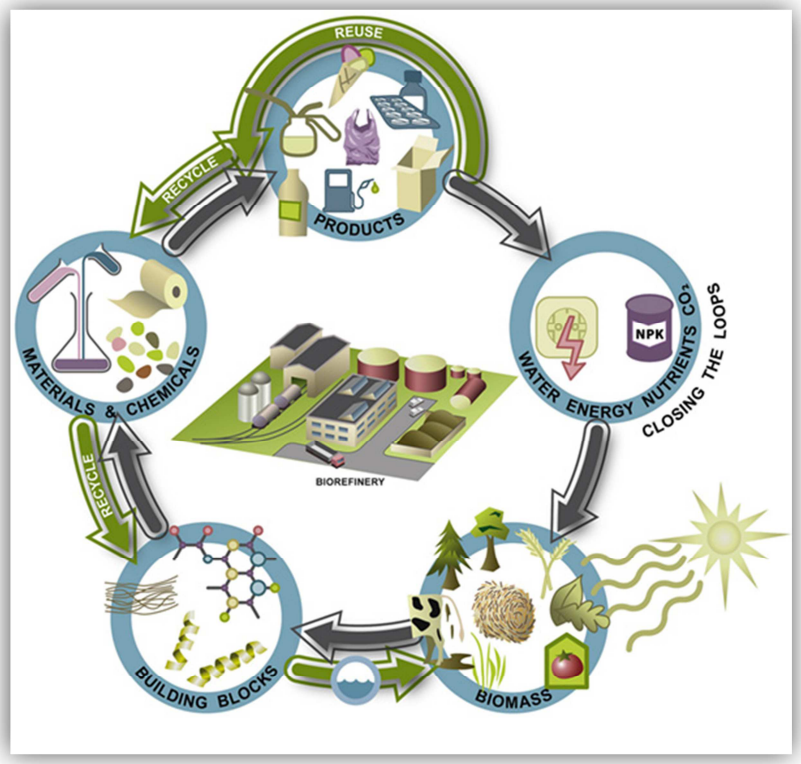


Figure 3. The Biorefinery concept. ©*Creative Commons Attribution 3.0*

1.4 Aims of the thesis

The overall goal of this PhD project is the valorization of selected lignocellulosic biomasses as source of fermentable sugars and novel biocatalysts (cellulases, hemicellulases and accessory enzymes) for the production of bioethanol and high value added products.
The core of this aim is:

1. the production of fermentable sugars from lignocellulose by:
    a. the valorization of the perennial biomass crop *Arundo donax*;
    b. the upgrading of newspaper fraction of Municipal Solid Waste (MSW);
2. the identification of novel lignocellulose-degrading biocatalysts (cellulases, hemicellulases and accessory enzymes) by:
    a. the traditional microorganisms cultivation based approach;
    b. the metagenomic approach;
3. the development of a *category II* Consolidated BioProcessing (CBP) system to produce bioethanol in one-step way combining cellulase production, enzymatic hydrolysis and fermentation.

As far as the production of fermentable sugars from lignocellulose is concerned, the strategy was the evaluation of two different lignocellulosic biomasses (*Arundo donax* and Newspaper Waste) as source of monosaccharides, by enzymatic hydrolysis after Ammonia Fiber Expansion (AFEX) or Extractive Ammonia (EA) pretreatment. The ability of the recombinant endocellulase rCelStrep, α-L-arabinofuranosidase rPoAbf and its evolved variant rPoAbf F435Y/Y446F to improve the saccharification yields was evaluated. All the research activities related to this objective were carried out at the Department of Chemical Engineering and Materials

Science, DOE Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, MI, USA.

Regarding the identification of novel (hemi)cellulases and accessory enzymes, three dedicated lignocellulosic energy crops, *Arundo donax, Eucalyptus camaldulensis* and *Populus nigra* were subjected to biodegradation under natural conditions in field or underwood in order to enhance the microbial diversity. A traditional microorganisms cultivation based approach was carried out to isolate and select new well-adapted strains in order to identify and characterize novel (hemi)cellulolytic enzymes produced by the selected microorganisms. Moreover, the DNAs from microorganisms adherent to the plant biomasses were subjected to metagenomic approach by both function-driven analysis of an expression library and sequence-driven analysis in order to identify putative genes and enzymes involved in lignocellulose degradation.

As far as the development of a one-step Consolidated BioProcessing of lignocelluloses to bioethanol (Category II CBP) is concerned, a strategy to engineer a *Saccharomyces cerevisiae* industrial strain with the cDNA sequence encoding for the endocellulase CelStrep was tested.

Chapter 2

Lignocellulosic biomasses as source of fermentable sugars

As previously reported, considering mainly the output/input energy ratio, availability and low cost, lignocellulosic material represents a promising option as feedstock to produce monosaccharides useful for the further production of valuable bioproducts via fermentation. Nowadays the research is moving towards the evaluation of different biomasses, such as agro-industrial wastes, crops on marginal lands and fractions of Municipal Solid Waste, as raw materials in the conversion processes. For this purpose, in this chapter one perennial biomass crop (*Arundo donax*) able to grow on marginal and/or polluted areas and one fraction of Municipal Solid Waste (Newspaper) were tested as sources of fermentable sugars.

All the research activities described below were carried out at the Department of Chemical Engineering and Materials Science, DOE Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, MI, USA with the supervision of prof. Venkatesh Balan.

2.1 – Valorization of crops on marginal lands. The case of *Arundo donax*

Among the different perennial biomass crops, one of the most interesting is giant reed (*Arundo donax L.*), because of its tolerance to a wide range of environmental stresses (such as high soil salinity, periods of severe drought), its ability to grow in several types of soils (ranging from heavy clay to loose sands and gravelly soil) and on degraded or contaminated lands (Fagnano et al., 2015). Moreover, *Arundo donax* biomass consists mainly of cellulose and hemicellulose with a low level of lignin, thus proving to be very interesting for biofuels and biochemicals production.

The following paper **"Assessment of bacterial and fungal (hemi)cellulose-degrading enzymes in saccharification of ammonia fibre expansion-pretreated *Arundo donax*" (Paper I)** is the first manuscript aimed at evaluating the bioconversion of *Arundo donax* biomass into monosaccharides after Ammonia Fiber Expansion (AFEX) pretreatment. This work was supported by a grant from the Ministero dell'Università e della Ricerca Scientifica — Industrial Research Project "Integrated agro-industrial chains with high energy efficiency for the development of eco-compatible processes of energy and biochemical production from renewable sources and for the land valorization (EnerbioChem)" PON01_01966, funded in the frame of Operative National Programme Research and Competitiveness 2007–2013 D. D. Prot. n. 01/Ric. 18.1.2010.

Appl Microbiol Biotechnol (2016) 100:2213–2224
DOI 10.1007/s00253-015-7066-3

Assessment of bacterial and fungal (hemi)cellulose-degrading enzymes in saccharification of ammonia fibre expansion-pretreated *Arundo donax*

Simona Giacobbe[1] · Venkatesh Balan[2] · Salvatore Montella[1] · Massimo Fagnano[3] · Mauro Mori[3] · Vincenza Faraco[1]

Received: 2 July 2015 / Revised: 1 October 2015 / Accepted: 7 October 2015 / Published online: 31 October 2015

Abstract This study reports enzymatic hydrolysis of the biomass of the giant reed (*Arundo donax* L.) after ammonia fibre expansion (AFEX) pretreatment. In particular, the capacity of the arabinofuranosidase from the fungus *Pleurotus ostreatus* recombinantly expressed in *Pichia pastoris* rPoAbf, its evolved mutant rPoAbf F435Y/Y446F and the endocellulase from *Streptomyces* sp. *G12* CelStrep recombinantly expressed in *Escherichia coli* to enhance the hydrolysis of AFEX-treated *A. donax* was investigated, using the corn stover as reference feedstock. The investigated enzymes were assayed using a mixture of purified cellulases (CBHI, CBHII, EGI and βG), endoxylanases (LX3, LX4) and accessory hemicellulases (LarbF and LβX) as reference enzyme mixture and substituting EGI with rCelStrep and LarbF with rPoAbf or rPoAbf F435Y/Y446F. The use of rPoAbf F435Y/Y446F in the substitution of LarbF led to improvements in sugar conversion, giving a glucan, xylan and arabinan conversion after 72 h of around 62, 63 and 80 %, respectively, similar or higher than those (44, 66 and 55 %) achieved by 72 h hydrolysis with commercial enzymes Novozymes Cellic®, Ctec3 and Htec3. The enzymes rPoAbf, rPoAbf F435Y/Y446F and rCelStrep were also investigated for their effect on hydrolysis of AFEX-pretreated *A. donax* by addition to commercial enzyme mixture Novozymes Cellic®, Ctec3 and Htec3, and it was shown that the addition of rPoAbf and its evolved mutant rPoAbf F435Y/Y446F enhanced both xylan and arabinan conversions, which achieved 80 % after 6 days of saccharification with rPoAbf F435Y/Y446F.

Keywords Lignocellulose · Pretreatment · Cellulase · Arabinofuranosidase

✉ Vincenza Faraco
vfaraco@unina.it

[1] Department of Chemical Sciences, University of Naples "Federico II", Complesso Universitario Monte S. Angelo, via Cintia, Napoli, Italy

[2] Department of Chemical Engineering and Materials Science, DOE Great Lakes Bioenergy Research Center, Michigan State University, Lansing, MI 48823, USA

[3] Department of Agriculture, University of Naples "Federico II", Portici, Napoli, Italy

Introduction

The policies for climate change mitigation promote the replacement of fossil fuels and petroleum-based products with alternative bioproducts from renewable resources such as biomass crops (Kajaste 2014). However, the incentives given to produce energy crops could have devastating effects on agricultural markets, as several food grain-producing lands will be taken away, reducing food production and consequently increasing food prices (Scheidel and Sorman 2012).

To avoid any competition for land between food and non-food crops, several efforts are under way to use crop lands not suitable for the traditional food crops for growing dedicated energy crops.

In Mediterranean environments, hilly areas are considered not appropriate for the traditional cereal production because yield and gross income are very low and because the traditional cropping system (deep soil tillage at the end of August and sowing in November) causes extreme vulnerability to soil erosion (Diodato et al. 2009, 2011; Fagnano et al. 2012). In these conditions, perennial biomass crops such as the giant reed (*Arundo donax* L.) proved to reduce soil erosion and to increase the potential gross income of farmers (Fagnano et al.

2015), with favourable environmental impacts (Forte et al. 2015).

The giant reed (*A. donax* L.) is also suggested for other areas not suitable for food crops such as polluted soils since it acts as a phyto-remediating agent and it also allows the production of large amounts of cellulose and hemicellulose (Fiorentino et al. 2010, 2013).

Many studies have shown that sugar polymers (cellulose and hemicellulose) present in biomasses can be hydrolyzed into fermentable sugars and then converted into fuels and chemicals (Wettstein et al. 2012; Kobayashi and Fukuoka 2013; Kajaste 2014). The Pacific Northwest National Laboratory (PNNL) and National Renewable Energy Laboratory (NREL) identified 12 building blocks produced through sugar conversions which could be used to obtain a variety of high-value biobased products. These include succinic acid, glucaric acid, aspartic acid, glycerol, sorbitol and xylitol/arabinitol (Hermann et al. 2007).

The production of bioproducts from lignocellulosic biomass requires several steps (Kumar et al. 2009), including an initial pretreatment step needed to break the lignin barrier and make cellulose and hemicellulose accessible to the enzymes during the hydrolysis to produce fermentable sugars. Enzymatic hydrolysis of sugar polymers using biomass-degrading enzymes such as cellulase and hemicellulase is preferred over dilute acid hydrolysis because of the higher conversion and lower environmental impact (Taherzadeh and Keikhosro 2007). Due to the complexity of the polysaccharides of the pretreated biomass, a tailor-made enzyme cocktail is required to achieve a more efficient saccharification (Gao et al. 2010a, 2011). Besides the enzymes involved in the hydrolysis of cellulose and the backbone of hemicellulose, some other enzymes with hydrolytic ability towards glycoside branches in hemicellulose were proven to improve the overall monosaccharide yield.

This study was aimed at evaluating the conversion of *A. donax* into monosaccharides for bioethanol or bioproducts production by enzymatic hydrolysis of the polysaccharides issued from pretreatment of the biomass by ammonia fibre expansion (AFEX). To the best of our knowledge, this is the first manuscript on saccharification of AFEX-pretreated *A. donax*.

In particular, the capacity of the cellulase CelStrep from *Streptomyces* sp. G12 recombinantly expressed in *Escherichia coli* (Amore et al. 2012b) and α-L-arabinofuranosidase PoAbf from the fungus *Pleurotus ostreatus* recombinantly expressed in *Pichia pastoris* (Amore et al. 2012a) and its evolved mutant rPoAbf F435Y/Y446F (Giacobbe et al. 2014) to improve the sugar yields of AFEX-pretreated *A. donax* was investigated. The rPoAbf F435Y/Y446F mutant had been previously developed with the aim to improve the catalytic efficiency of the investigated α-L-arabinofuranosidase (Giacobbe et al. 2014). It was further investigated in biomass conversion due to its ability to hydrolyze both soluble and insoluble substrates better than that of the rPoAbf wild type. Differently from our previous work (Marcolongo et al. 2014), the investigated enzymes were used in combination with purified fungal and bacterial enzymes to define a tailor-made enzyme cocktail for *A. donax* hydrolysis in order to reduce the cost of the enzymatic saccharification process.

Materials and methods

Feedstock

Biomasses used in this study for pretreatment and saccharification experiments include corn stover and the giant reed (*A. donax* L.). *A. donax* was produced in marginal lands of southern Italy with a low-input cropping system (Forte et al. 2015), and corn stover was provided by Michigan State University (MSU). The biomasses were milled with a 2-mm-diameter sieve and stored under dry conditions at room temperature until further use. The moisture content was measured using a moisture analyser (Sartorius MA35M, Elk Grove, IL).

Compositional analysis

Compositional analyses of *A. donax* and corn stover biomasses were performed following the National Renewable Energy Laboratory (NREL) Laboratory Analytical Procedures (LAPs) standard protocols:

- "Preparation of samples for compositional analysis" (Hames et al. 2008)
- "Determination of structural carbohydrates and lignin in Biomass"
- "Determination of Total Solids in Biomass and Total Dissolved Solids in Liquid Process Samples" (Sluiter et al. 2008)
- "Determination of Ash in Biomass" (Sluiter et al. 2005b)
- "Determination of structural carbohydrates and lignin in biomass" (Sluiter et al. 2005a)

Monomeric sugars were quantified using a Bio-Rad Aminex HPX-87H high-performance liquid chromatography (HPLC) column using 5 mM sulphuric acid as mobile phase.

The variability between experiments is reflected by the standard deviation reported.

AFEX pretreatment

Corn stover and *A. donax* were pretreated through the AFEX method by varying the ammonia to biomass ratio (1:1 and 2:1), reaction temperature (100–160 °C), moisture (60–

233 % on dry weight basis) and fixed residence time (15 min). AFEX was carried out in a high-pressure stainless steel vessel. Biomass was first loaded into the vessel with appropriate moisture after taking into consideration the moisture content of original biomass. Then, the reactor was closed and vacuum applied to remove residual air in the reactor. The required amount of liquid ammonia was loaded into the reactor using an ammonia delivery pump. The vessel was heated by an external mantle and the biomass was mixed during the AFEX process for 15 min. As the temperature of the reactor increased, the pressure in the vessel rose (between 200 and 400 psi) depending on the ammonia to biomass loading. The pressure was released from the vessel and ammonia was vented in a fume hood. The pretreated biomass was moved to a tray and dried in a fume hood overnight to remove residual ammonia. Then, dry AFEX-treated biomass was stored in a sealed polythene bag at 4 °C until further use.

Purified enzymes and their source

The following purified enzymes were tested in this study:

- Fungal cellobiohydrolase I (CBH I; glycoside hydrolase (GH) family 7A), cellobiohydrolase II (CBH II; GH family 6A) and endoglucanase I (EG I; GH family 7B) were purified from Spezyme CP using several chromatography steps (size-exclusion chromatography, anion and cation exchange chromatography, hydrophobic interaction chromatography and affinity chromatography) as described in Gao et al. (2010a). The enzymes were used at a concentration of 3.32 mg/g glucan each.
- Fungal β-glucosidase (βG; GH family 3) was purified from Novozyme 188 using anion exchange chromatography followed by cation exchange chromatography as described in Gao et al. (2010a) and used at a concentration of 2 mg/g glucan.
- Bacterial xylanases (LX3, GH family 10; LX4, GH family 11) from *Clostridium thermocellum* library (Gao et al. 2011) (bacterial cell concentrate was a kind gift of Dr. Paul Weimer, United States Department of Agriculture, Agricultural Research Service, United States Dairy Forage Research Center, WI, USA) were recombinantly expressed in *E. coli* and purified using HIS-select nickel affinity chromatography. The enzymes were used at a concentration of 1.66 mg/g glucan each.
- Bacterial β-xylosidase (LβX; GH family 52) from *Geobacillus stearothermophilus* XYNB2 (Gao et al. 2010b) was recombinantly expressed in *E. coli* and purified using HIS-select nickel affinity chromatography. The enzyme was used at 0.6 mg/g glucan.
- Bacterial α-arabinofuranosidase (LarbF, GH family 51) from *Geobacillus* sp. G11MC16 was recombinantly expressed in *E. coli* and purified using HIS-select nickel affinity chromatography. The enzyme was loaded at 0.6 mg/g glucan (Gao et al. 2011).
- Bacterial CelStrep (EMBL accession number HE862416) from *Streptomyces* sp. G12 was recombinantly expressed in *E. coli* (Amore et al. 2012b). The enzyme was purified by ammonium sulphate precipitation followed by hydrophobic interaction chromatography, as described in Amore et al. (2012b) and used at 3.32 mg/g glucan.
- Fungal α-L-arabinofuranosidase (rPoAbf) from *P. ostreatus* (ATCC 66376) was recombinantly expressed in *P. pastoris* (Amore et al. 2012a). The enzyme was purified by ammonium sulphate precipitation followed by hydrophobic interaction chromatography, as described in Amore et al. (2012a) and loaded at 0.6 mg/g glucan.
- The variant F435Y/Y446F of PoAbf, previously selected from a directed evolution library of rPoAbf, was recombinantly expressed in *P. pastoris* (Giacobbe et al. 2014), purified as rPoAbf wild type and loaded at 0.6 mg/g glucan.

Determination of protein concentration

Purified protein concentration was determined by the Pierce (Pierce Biotechnology, Rockford, USA) bicinchoninic acid (BCA) assay kit following the manufacturer's instructions. Bovine serum albumin (BSA) was used as standard.

Enzymatic hydrolysis

The bioconversion experiments were performed in five vials at 1 % (*w*/*w*) glucan loading in 50 mM citrate buffer (pH 4.8) with the desired enzymes. 0.5 mM sodium azide was used to prevent microbial and fungal growth. The saccharification was performed at 50 °C and 250 rpm in a shaking incubator. Sampling was done every 24 h to evaluate the sugar composition using high-performance liquid chromatography (HPLC) system. The commercial enzyme mixture Novozymes Cellic® (60 % Ctec3 and 40 % Htec3) was used (at a loading of 15 mg/g of glucan) for saccharification experiments to evaluate the best pretreatment conditions. The hydrolysis of AFEX-treated biomasses was performed with the following enzymes at a loading of around 16.5 mg/g glucan (Table 1). An enzymatic mixture named mix A was prepared, including CBH I, cellobiohydrolase II, EG I at a concentration of 3.32 mg/g glucan each, βG 2 mg/g glucan, xylanases (LX3, LX4) 1.66 mg/g glucan each and LβX and LarbF 0.6 mg/g glucan each. In mix B, the EG I was replaced by cellulase rCelStrep; in mix C, the LarbF was replaced by the rPoAbf; in mix D, the LarbF was replaced by a mutant of rPoAbf named rPoAbf F435Y/Y446F; in mix E, both EGI and LarbF were replaced by rCelStrep and rPoAbf; in mix F, both EGI and LarbF were replaced by rCelStrep and rPoAbf F435Y/Y446F.

Table 1 Mixtures of bacterial and fungal enzymes tested on AFEX-treated biomasses

Individual enzyme loading (mg/g glucan)											
Mix	CBHI	CBHII	βG-	LX3	LX4	LβX	EGI	LarbF	rCelstrep	rPoAbf	rPoAbf F435Y/Y446F
A	3.32	3.32	2	1.66	1.66	0.6	3.32	0.6			
B	3.32	3.32	2	1.66	1.66	0.6		0.6	3.32		
C	3.32	3.32	2	1.66	1.66	0.6	3.32			0.6	
D	3.32	3.32	2	1.66	1.66	0.6	3.32				0.6
E	3.32	3.32	2	1.66	1.66	0.6		0.6	3.32	0.6	
F	3.32	3.32	2	1.66	1.66	0.6		0.6	3.32		0.6

Moreover, rCelStrep (3.32 mg/g glucan), rPoAbf (0.6 mg/g glucan) and rPoAbf mutant (0.6 mg/g glucan) were added to the commercial mixture Novozymes Cellic®.

Sugar analysis

About 200 μl hydrolysate collected at different times during enzyme hydrolysis was transferred to a centrifuge tube, heated to 100 °C for 10 min (to denature the enzymes), centrifuged at 8000 rpm for 10 min to remove the precipitates and then stored in a HPLC vial at −20 °C until further use. A Bio-Rad Aminex HPX-87P HPLC column was used to determine the monomeric sugar concentrations in the hydrolysate. All experiments were performed in triplicate. Shimadzu HPLC Prominence system (Columbia, MD, USA) with a refractive index detector (RID), was employed for analysing the sugars. Water was the mobile phase at a fixed flow rate of 0.6 ml/min, with isocratic elution. The column temperature was maintained at 60 °C, and the HPLC sample injection volume was 20 μl. Standard curves were generated using different concentrations of mixed sugars. A guard column with similar packing was used throughout the chromatography experiments.

Statistical analyses

One-way ANOVA followed by Tukey's HSD post hoc for pairwise comparison of means (at $P \leq 0.05$) was used to assess the difference in the sugar conversion of the different enzyme mixtures used to hydrolyze *A. donax* or corn stover biomass. Statistical analyses were performed using SPSS 13.0 statistical software package (SPSS Inc., Cary, NC, USA).

Results

Characterization of biomasses

Analyses of the macromolecular composition of the unpretreated biomasses *A. donax* and corn stover were performed, and the results, reported in Table 2, revealed that *A. donax* contains 61.41 % of structural polysaccharides, 26.24 % Klason lignin and 4.9 % ash, while corn stover contains 63.23 % of structural polysaccharides, 20.06 % Klason lignin and 6.17 % ash. Both the analysed biomasses contain significant percentages of C5 sugars mainly represented by xylan (~20 %) while arabinan represents only around 2 % of total dry weight and the C6 sugars consist mainly of glucan (>37 %).

Selection of AFEX pretreatment conditions

The yield of monosaccharides released during hydrolysis depends on various AFEX parameters such as moisture content, ammonia to biomass loading, residence time and temperature. Therefore, in order to assess the best pretreatment conditions for *A. donax*, three different AFEX treatments were carried out (Table 3). The AFEX-tested conditions, condition 1, condition 2 and condition 3 reported in Table 3, were chosen since they had been previously identified as the best conditions for the AFEX pretreatment of *Miscanthus* x *giganteus* (Murnen et al. 2007), corn stover (Balan et al. 2009b) and switchgrass (Bals et al. 2010), respectively. Since corn stover was used as the

Table 2 *A. donax* and corn stover composition before AFEX pretreatment (standard deviation obtained from three independent experiments)

Composition	*A. donax* (%, dry weight basis)	Corn stover (%, dry weight basis)
Moisture content	6.0±0.5	8.7±0.6
Ash	4.9±0.6	6.7±1.0
Structural carbohydrate		
Glucan	38.0±0.0	38.41±0.1
Xylan	21.0±0.1	19.9±1.0
Galactan	1±0.1	1.4±0.1
Arabinan	1.5±0.0	2.5±0.2
Mannan	N.D.	N.D.
Lignin		
Acid-insoluble lignin	26.2±0.1	20.1±0.7
Acid-soluble lignin	1.9±0.0	2.0±0.3

ND Not determined

Table 3 Experimental AFEX pretreatment conditions tested in this study and percentage of glucan and xylan conversion after 24, 72 and 168 h hydrolysis of *A. donax* and corn stover by the commercial enzymatic preparation consisting of Novozymes Cellic® 60 % Ctec3 and 40 % Htec3 (15 mg/g glucan)

Arundo donax									
AFEX conditions	Ammonia loading (kg/kg dry biomass)	Temperature (°C)	Moisture content (%)	Time (min)	Glucan conversion (%)			Xylan conversion (%)		
					24 h hydrolysis	72 h hydrolysis	168 h hydrolysis	24 h hydrolysis	72 h hydrolysis	168 h hydrolysis
1	1:1	100	80	15	34.4	41.2	49.7	55.6	60.0	66.9
2	1:1	130	60	15	39.0	44.1	56.8	66.4	65.9	75.5
3	2:1	160	233 (water soak)	15	22.4	24.0	28.3	22.4	24.5	29.6
Corn stover										
	Ammonia loading (kg/kg dry biomass)	Temperature (°C)	Moisture content (%)	Time (min)	Glucan conversion (%)			Xylan conversion (%)		
					24 h hydrolysis	72 h hydrolysis	168 h hydrolysis	24 h hydrolysis	72 h hydrolysis	168 h hydrolysis
	1:1	130	60	15	71.9	79.1	90.4	70.0	90.0	100.0

reference biomass in this study, the best AFEX condition for pretreatment of this biomass was tested to pretreat *A. donax*. Nevertheless, since *M.* x *giganteus* and switchgrass are perennial crops like *A. donax*, the best conditions for their AFEX pretreatment were also tested. The samples of *A. donax* pretreated in the different conditions were then enzymatically hydrolyzed for 168 h using a commercial enzymatic preparation consisting of Novozymes Cellic®, 60 % Ctec3 and 40 % Htec3. In Table 3, the glucan and xylan conversions obtained after 24, 72 and 168 h of enzymatic hydrolysis of *A. donax* after the three different AFEX pretreatments are reported in comparison to corn stover. Both glucan and xylan hydrolyses reached their maximum level with AFEX pretreatment condition 2: 130 °C, 1 kg ammonia/1 kg dry biomass and 60 % of moisture content. Based on these data, this condition was selected for further hydrolysis experiments.

Saccharification of *A. donax* and corn stover using purified enzyme cocktail and commercial enzyme cocktail

The enzymatic hydrolysis of biomass using commercial enzymes is not appropriate to evaluate the role of individual enzymes involved in the process, and it is difficult to obtain a tailor-made enzyme cocktail for the biomass of interest. Previous works have shown that an optimal cocktail of cellulolytic enzymes (CBH I, CBH II, EGI, βG) and a set of hemicellulases and accessory enzymes (LX3, LX4, LarbF, LβX) are required to improve the hydrolysis of AFEX-treated corn stover (Gao et al. 2010a, 2011). Based on these results, the enzyme mix A (CBH I, CBH II, EGI, βG, LX3, LX4, LarbF, LβX), previously optimized for corn stover (Gao et al. 2011), was used as reference enzymatic mixture to hydrolyze *A. donax*. Figure 1 shows the glucan, xylan and arabinan conversion percentages from corn stover and *A. donax* after AFEX treatment in selected condition 2, by 24- and 72-h-long hydrolyses with mix A.

As far as *A. donax* is concerned, although the enzyme mix A gave similar xylan hydrolysis (~63 % after 72 h) compared with the commercial enzymatic preparation, Novozymes Cellic®, 60 % Ctec3 and 40 % Htec3, reported in Table 2, lower glucan hydrolysis was achieved (less than 30 % after 72 h). As regards corn stover, mix A gave sugar conversions (Fig. 1) similar to those obtained with the Novozyme Cellic® enzyme cocktail (Table 3).

Effect of the enzymes rCelStrep, rPoAbf and its variant on saccharification of *A. donax* and corn stover in comparison with the enzymes EGI or LarbF

The effect of the arabinofuranosidase from the fungus *P. ostreatus* recombinantly expressed in *P. pastoris* rPoAbf, its evolved mutant expressed in the same yeast rPoAbf

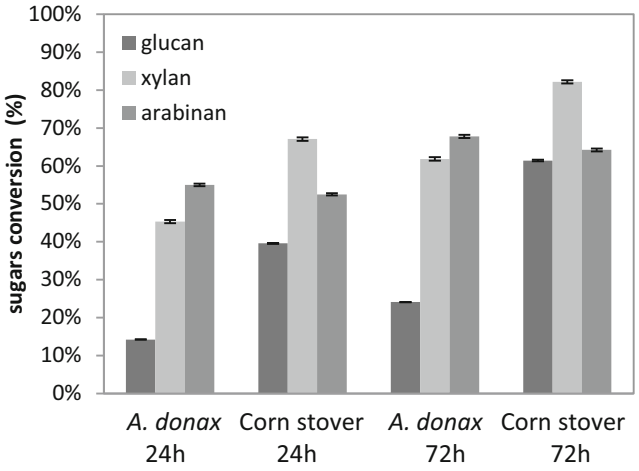


Fig. 1 Glucan, xylan and arabinan conversion after 24 and 72 h of hydrolysis on AFEX-treated corn stover and *A. donax* using the enzyme mix A containing the cellulases CBH I, CBH II and EG I loaded at 3.32 mg/g glucan each, βG loaded at 2 mg/g glucan, and the endoxylanases LX3 and LX4 and accessory hemicellulases LarbF and LβX loaded at 1.66 and 0.6 mg/g glucan each, respectively

F435Y/Y446F and the cellulase from *Streptomyces* sp. *G12* rCelStrep recombinantly expressed in *E. coli* in bioconversion of AFEX-pretreated *A. donax* and corn stover was tested. In these experiments, the enzyme mix A, containing cellulases (CBHI, CBHII, EGI and βG), endoxylanases (LX3, LX4) and accessory hemicellulases (LarbF and LβX), was used as reference enzyme mixture, and the effect of substituting EGI with rCelStrep and LarbF with rPoAbf or rPoAbf F435Y/Y446F was assessed. In more detail, the bacterial cellulase rCelStrep was used instead of the fungal endoglucanase EGI in the enzyme mix B while the fungal arabinofuranosidases rPoAbf was used instead of the bacterial arabinofuranosidase LarbF in the enzyme mix C. The evolved variant of rPoAbf (rPoAbf F435Y/Y446F) was used instead of the bacterial arabinofuranosidase LarbF in the enzyme mix D due to its ability to hydrolyze both soluble and insoluble substrates better than that of the rPoAbf wild type. Moreover, rCelStrep was also tested in combination with rPoAbf (mix E) or its variant (mix F) to check a possible synergism between bacterial and fungal enzymes.

Table 4 summarizes the conversion results of AFEX-pretreated *A. donax* biomass using different purified enzymatic mixtures. The sugar conversion results showed that the substitution of EGI with rCelstrep (mix B) gave significantly higher glucan conversion (28 %) after 72 h than that obtained with enzyme mix A (24.1 %), corresponding to 15 % increase in sugar conversion. On the other hand, the use of rCelstrep (mix B) decreased arabinan hydrolysis, giving an arabinan conversion 17 % lower than that obtained with the enzyme mix A after 72 h.

When the rPoAbf wild type was used instead of LarbF (mix C), no significant differences in glucan and xylan hydrolyses were observed. The use of rPoAbf wild type gave an arabinan hydrolysis 14 % significantly lower than that obtained with mix A after 72 h hydrolysis.

The use of rPoAbf F435Y/Y446F instead of LarbF (enzyme mix D in comparison with enzyme mix A) and the wild-type rPoAbf (mix D in comparison with mix C) led to improvements in sugar conversion. The use of mix D gave 60 % glucan conversion after 72 h, which is 2.5 times higher than that reached with mix A and mix C. The use of mix D gave 53 % xylan conversion after 24 h, which is 16 % higher than that obtained with mix A. Moreover, the xylan conversion after 72 h saccharification with mix D achieved 63 %, similar to that obtained with mix A and mix C. The arabinan conversion was found to be 78 % after 72 h, which is 15 and 26 % higher than that obtained with the mix A and mix C, respectively (Fig. 2).

The effect of rCelstrep in combination with rPoAbf (mix E) and rPoAbf F435Y/Y446F (mix F) was also analysed (Table 4). After 24 h, glucan conversion was found to be not significantly higher than mix A, for both mix E and mix F. However, about 20 % glucan conversion was obtained after 72 h for enzyme mix E, which is 14.5 % lower than that obtained with enzyme mix A; also, the xylan conversion resulted significantly lower than that with enzyme mix A after 72 h hydrolysis. As regards mix F, no significant differences with enzyme mix A were found in xylan hydrolysis after both 24 and 72 h. Mix E gave 20 and 28 % arabinan conversion, which are 64 and 56 % lower than that with enzyme mix A, after 24 and 72 h, respectively. On the other hand, no significant differences in arabinan conversion were obtained with mix F in comparison to mix A.

Hydrolysis of AFEX-treated corn stover using purified enzyme cocktail mixtures showed that no significant increase in sugar conversion was achieved replacing enzyme mix A with mix B, mix E and mix F (Table 5). On the other hand, enzyme mix C and mix D gave 8 % higher glucan conversion after 24 h when compared to enzyme mix A. However, the sugar conversion decreased after 72 h hydrolysis. Enzyme mix C and mix D gave 57.4 and 63.2 % arabinan conversion, respectively, after 24 h, which is 8.5 and 17 % higher than that with mix A.

It is worth noting that the evolved variant rPoAbf F435Y7Y446F gave a better arabinan conversion than the wild-type enzyme for both the biomasses.

Synergy between commercial enzyme cocktail and purified enzymes (rPoAbf and/or rCelstrep)

To study synergy between the commercial enzyme cocktail Novozymes Cellic®, Ctec3 and Htec3, and the purified enzymes rPoAbf, its evolved mutant rPoAbf F435Y/Y446F and rCelstrep, hydrolysis of AFEX-pretreated *A. donax* and corn stover was carried out adding each tested enzyme or their

Table 4 *A. donax* sugar conversion in 50 mM citrate buffer pH 4.5 and *T* 50 °C, for 24 and 72 h. The values represent the means±SD of three replicates. Different superscript letters after the values indicate significant differences ($P \leq 0.05$) on the same column within samples collected after 24 or 72 h

Enzyme mixture	Purified enzyme cocktail	Glucan conversion (%)	Xylan conversion (%)	Arabinan conversion (%)
24 h hydrolysis				
Mix A	CBHI-CBHII-βG-LX3-LX4-LβX+EGI & LarbF	14.2±1.6^{ab}	45.4±0.5$^{a–c}$	55.0±2.5^{b}
Mix B	CBHI-CBHII-βG-LX3-LX4-LβX+rCelstrep & LarbF	11.8±2.9^{a}	43.7±2.5^{ab}	52.3±2.1^{b}
Mix C	CBHI-CBHII-βG-LX3-LX4-LβX+EGI & rPoAbf	16.9±0.2^{bc}	49.2±0.5^{cd}	56.1±4.0^{b}
Mix D	CBHI-CBHII-βG-LX3-LX4-LβX+EGI & rPoAbf F435Y/Y446F	17.5±1.8^{c}	52.5±4.2^{d}	64.9±6.1^{c}
Mix E	CBHI-CBHII-βG-LX3-LX4-LβX+rCelstrep & rPoAbf	16.6±0.3^{bc}	41.8±0.1^{a}	19.9±0.3^{a}
Mix F	CBHI-CBHII-βG-LX3-LX4-LβX+rCelstrep & rPoAbf F435Y/Y446F	16.4±0.1^{bc}	46.1±0.8^{bc}	53.1±1.4^{b}
72 h hydrolysis				
Mix A	CBHI-CBHII-βG-LX3-LX4-LβX+EGI & LarbF	24.1±1.1^{b}	61.8±0.9^{b}	67.8±0.6^{c}
Mix B	CBHI-CBHII-βG-LX3-LX4-LβX+rCelstrep & LarbF	28.5±2.6^{c}	59.1±6.2^{b}	56.3±0.3^{b}
Mix C	CBHI-CBHII-βG-LX3-LX4-LβX+EGI & rPoAbf	24.6±1.5^{b}	61.5±0.9^{b}	57.8±2.6^{b}
Mix D	CBHI-CBHII-βG-LX3-LX4-LβX+EGI & rPoAbf F435Y/Y446F	61.6±0.4^{d}	62.7±4.1^{b}	78.2±8.2^{d}
Mix E	CBHI-CBHII-βG-LX3-LX4-LβX+rCelstrep & rPoAbf	20.6±1.2^{a}	48.5±0.7^{a}	27.9±0.7^{a}
Mix F	CBHI-CBHII-βG-LX3-LX4-LβX+rCelstrep & rPoAbf F435Y/Y446F	22.1±1.2^{ab}	57.5±1.4^{b}	60.9±4.0^{bc}

combinations (rCelstrep and rPoAbf or rCelstrep and rPoAbf F435Y/Y446F) to the Novozymes cocktail.

The glucan conversion using commercial enzyme cocktail was found to be higher when compared to supplementation with wild-type rPoAbf or its evolved variant or rCelstrep or both enzymes (Fig. 3a). These results could be explained by hypothesizing an anti-synergism between the activity of (hemi)cellulases in commercial mixture and purified ones. The negative level of synergism was already described for commercial enzymes involved in biomass degradation and purified ones (Woodward J 1991; Morrison 2014).

Supplementing rPoAbf and its evolved variant gave 70 % xylan conversion, 7 % higher than that obtained by the commercial preparation. After 6 days of hydrolysis, 80 % xylan conversion was achieved when rPoAbf F435Y/Y446F was used along with commercial enzymes (Fig. 3b). The arabinan conversion was also improved by supplementing rPoAbf or variant to the commercial enzyme cocktail by 27 and 36 %, respectively, after 24 h of hydrolysis (Fig. 3c). A maximum arabinan conversion of 80 % was achieved after 6 days of hydrolysis when rPoAbf was added to commercial enzymes.

In the case of corn stover, no appreciable increase in glucan conversion was observed when commercial preparation was supplemented with wild-type rPoAbf or its variant or rCelstrep or both enzymes (data not shown). However, 99 % xylan conversion was achieved after 24 h of hydrolysis with commercial enzymes supplemented with rPoAbf wild type (Fig. 4a). The same xylan conversion was achieved using commercial preparation and commercial preparation in combination with rPoAbf F435Y/Y446F or rCelstrep or both

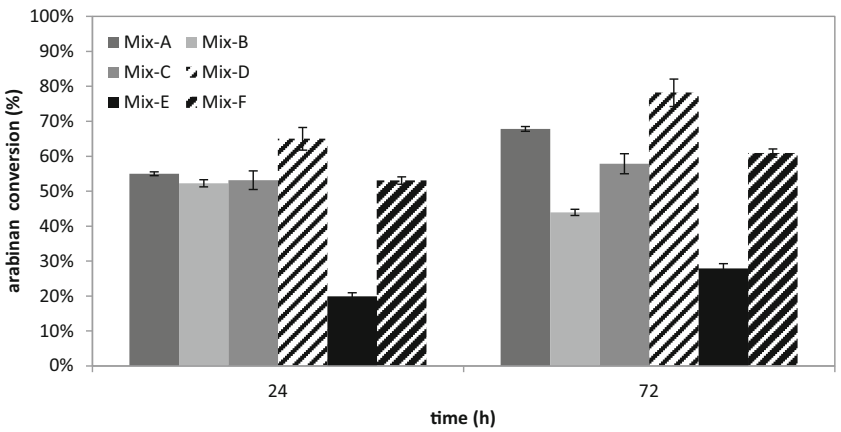


Fig. 2 Arabinan conversion after 24 and 72 h hydrolysis of ammonia fibre expansion (AFEX)-treated *A. donax* with the enzymatic mixtures reported in Table 4

Table 5 Corn stover sugar conversion in 50 mM citrate buffer pH 4.5, *T* 50 °C. The values represent the means±SD of three replicates. Different superscript letters after the values indicate significant differences ($P \leq 0.05$) on the same column within samples collected after 24 or 72 h

Enzyme mixture	Purified enzyme cocktail	Glucan conversion (%)	Xylan conversion (%)	Arabinan conversion (%)
24 h hydrolysis				
Mix A	CBHI-CBHII-βG-LX3-LX4-LβX+EGI & LarbF	39.6±0.8^{c}	67.1±1.2^{c}	52.5±1.3^{b}
Mix B	CBHI-CBHII-βG-LX3-LX4-LβX+rCelstrep & LarbF	32.7±0.1^{a}	67.5±0.1^{c}	53.5±0.1^{bc}
Mix C	CBHI-CBHII-βG-LX3-LX4-LβX+EGI & rPoAbf	42.8±0.7^{d}	72.8±1.1^{d}	57.4±0.8^{d}
Mix D	CBHI-CBHII-βG-LX3-LX4-LβX+EGI & rPoAbf F435Y/Y446F	42.7±0.2^{d}	73.6±0.4^{d}	63.2±0.2^{e}
Mix E	CBHI-CBHII-βG-LX3-LX4-LβX+rCelstrep & rPoAbf	38.3±1.1^{b}	59.6±1.2^{a}	54.1±2.0^{c}
Mix F	CBHI-CBHII-βG-LX3-LX4-LβX+rCelstrep & rPoAbf F435Y/Y446F	39.2±0.7^{c}	65.5±1.1^{b}	16.4±1.2^{a}
72 h hydrolysis				
Mix A	CBHI-CBHII-βG-LX3-LX4-LβX+EGI & LarbF	61.4±1.2^{f}	82.2±1.5^{e}	64.2±2.0^{e}
Mix B	CBHI-CBHII-βG -LX3-LX4-LβX+rCelstrep & LarbF	56.3±0.7^{e}	83.6±2.1^{e}	14.4±0.2^{a}
Mix C	CBHI-CBHII-βG -LX3-LX4-LβX+EGI & rPoAbf	11.7±1.0^{a}	77.0±1.2^{d}	56.0±0.1^{c}
Mix D	CBHI-CBHII-βG -LX3-LX4-LβX+EGI & rPoAbf F435Y/Y446F	41.2±0.6^{b}	73.4±0.4^{c}	63.5±0.2^{e}
Mix E	CBHI-CBHII-βG -LX3-LX4-LβX+rCelstrep & rPoAbf	47.2±0.8^{c}	57.9±0.2^{a}	59.2±1.4^{d}
Mix F	CBHI-CBHII-βG -LX3-LX4-LβX+rCelstrep & rPoAbf F435Y/Y446F	48.7±1.2^{d}	65.5±0.5^{b}	27.2±0.2^{b}

enzymes after 6 days of hydrolysis. Supplementation of commercial enzyme cocktail with rPoAbf gave 70 % of arabinan conversion after 24 h of hydrolysis (Fig. 4b).

Discussion

This study was focused on pretreatment and saccharification of the giant reed (*A. donax* L.) using corn stover as reference lignocellulosic biomass. *A. donax* is a biomass crop that can be cultivated with high yields in marginal areas not suitable for the traditional food crops, gaining at the same time several environmental benefits such as soil protection from erosion, C storage into the soil and phytoremediation of polluted soils (Fagnano et al. 2015; Forte et al. 2015).

The data of macromolecular composition of *A. donax* and corn stover used in this study was revealed to be in agreement with data reported in previous studies (Caparrós et al. 2007; Li et al. 2011; Scordia et al. 2012).

The pretreatment methods so far tested and reported in literature for *A. donax* include the steam explosion, acid-catalyzed steam explosion and ammonia soaking (De Bari et al. 2013; Van Den Brink et al. 2013; Marcolongo et al. 2014). However, most of these methods solubilize large portions of xylan from this biomass.

On the other hand, ammonia fibre expansion (AFEX) is a dry to dry process which preserves all the original carbohydrates (Gao et al. 2011 and 2013; Li et al. 2011; Harun et al. 2013; Uppugundla et al. 2014) without any loss. This pretreatment has never been used for *A. donax*. Therefore, this pretreatment process was selected for this study, in order to avoid the loss of the significant amounts of xylan present in the investigated biomass (Table 2).

Corn stover was chosen as reference lignocellulosic biomass because AFEX pretreatment had been previously carried out on it in several works giving high yields of both glucose and xylose (Balan et al. 2009; Li et al. 2011). It is worthy of note that the best AFEX condition selected in this study for the biomass of *A. donax* (condition 2: 130 °C, 1 kg ammonia/1 kg dry biomass and 60 % of moisture content) was the same pretreatment condition previously adopted for corn stover (Gao et al. 2011).

The effect of the α-L-arabinofuranosidase rPoAbf, its evolved mutant rPoAbf F435Y/Y446F and the cellulase rCelStrep, to enhance the saccharification of the pretreated giant reed (*A. donax* L.) in substitution of the purified enzymes LarbF or EGI, respectively, was tested in bioconversion of AFEX-pretreated *A. donax*. Moreover, synergy between commercial enzyme cocktail Novozymes Cellic®, Ctec3 and Htec3, and rPoAbf, rPoAbf F435Y/Y446F and rCelStrep in hydrolysis of AFEX-pretreated *A. donax* and corn stover was also investigated. Corn stover was the reference feedstock for all the saccharification experiments.

As the most positive effects observed for *A. donax* conversion, the use of the rPoAbf F435Y/Y446F within the mixture of purified enzymes allowed achieving better glucan, xylan and arabinan conversions than LarbF and rPoAbf wild type. In particular, rPoAbf F435Y/Y446F gave a glucan conversion after 72 h 2.5 times higher than that obtained with the other enzymes, a xylan conversion after 24 h 16 % higher than that

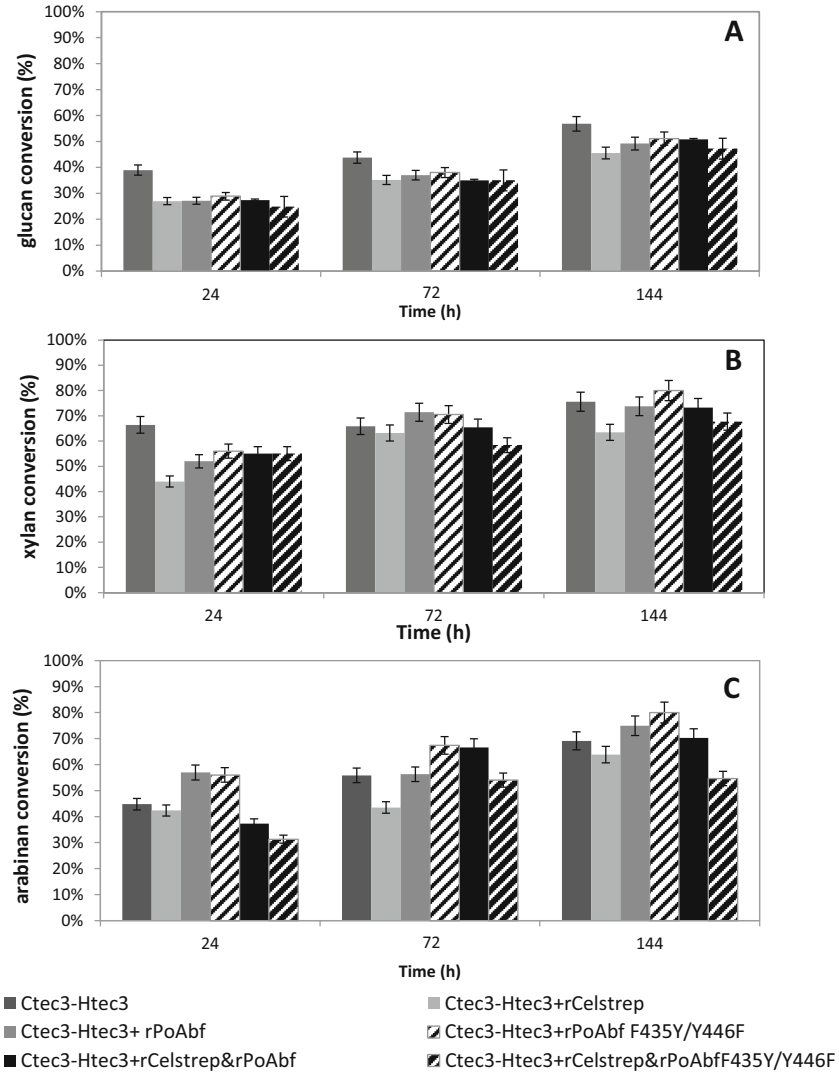


Fig. 3 Glucan (**a**), xylan (**b**) and arabinan (**c**) conversion of AFEX-treated *A. donax* achieved using commercial enzyme preparation Novozymes Cellic®, 60 % Ctec3 and 40 % Htec3, with or without addition of rPoabf, its variant F435Y/Y446F, rCelstrep or a combination of these enzymes

obtained with LarbF and an arabinan conversion after 72 h 15 and 26 % higher than that obtained with LarbF and the wild-type rPoAbf, respectively. The improvement of glucan conversion with the enzymatic cocktail containing rPoAbf F435Y/Y446F can be explained with the enhancement of the accessibility of this polysaccharide due to xylan hydrolysis, as previously reported (Polizeli et al. 2005).

The higher glucan conversion yield obtained substituting EGI with rCelstrep in the mixture of purified enzymes suggests a synergistic role of the bacterial cellulase rCelstrep when mixed with the fungal cellulases CBHI, CBHII and βG on glucan conversion. The use of Celstrep in combination with rPoAbf and its variant gave lower sugar conversion than those achieved using rPoAbf and its variant in combination with EGI, which could indicate a lower synergism between bacterial cellulase and fungal arabinofuranosidase than that between fungal cellulase and arabinofuranosidase. Moreover, this effect was more pronounced at longer times, which could be due to the loss of degree of synergy during the hydrolysis caused by structural modifications of the substrate as reported by Andersen et al. (2008) or saturation of sites for enzyme activity as proposed by Boisset et al. (2001).

As regards corn stover biomass, no significant improvement of sugar conversion was observed when the reference mixture of purified enzymes (mix A previously optimized for this biomass by Gao et al. 2011) was replaced by the other purified enzymes tested in this work.

The comparison of the effects of rPoAbf wild type and rPoAbf F435Y/Y446F in the presence of rCelstrep showed that the use of rPoAbf wild type led to the highest arabinan conversion from corn stover (mix E), while for *A. donax* biomass, the best conversion was obtained using rPoAbf F435Y/Y446F (mix F). These results could be explained by hypothesizing a diverse specificity and synergism of the enzymes towards the different biomasses.

As regards *A. donax* saccharification experiments performed with commercial enzyme cocktail Novozymes

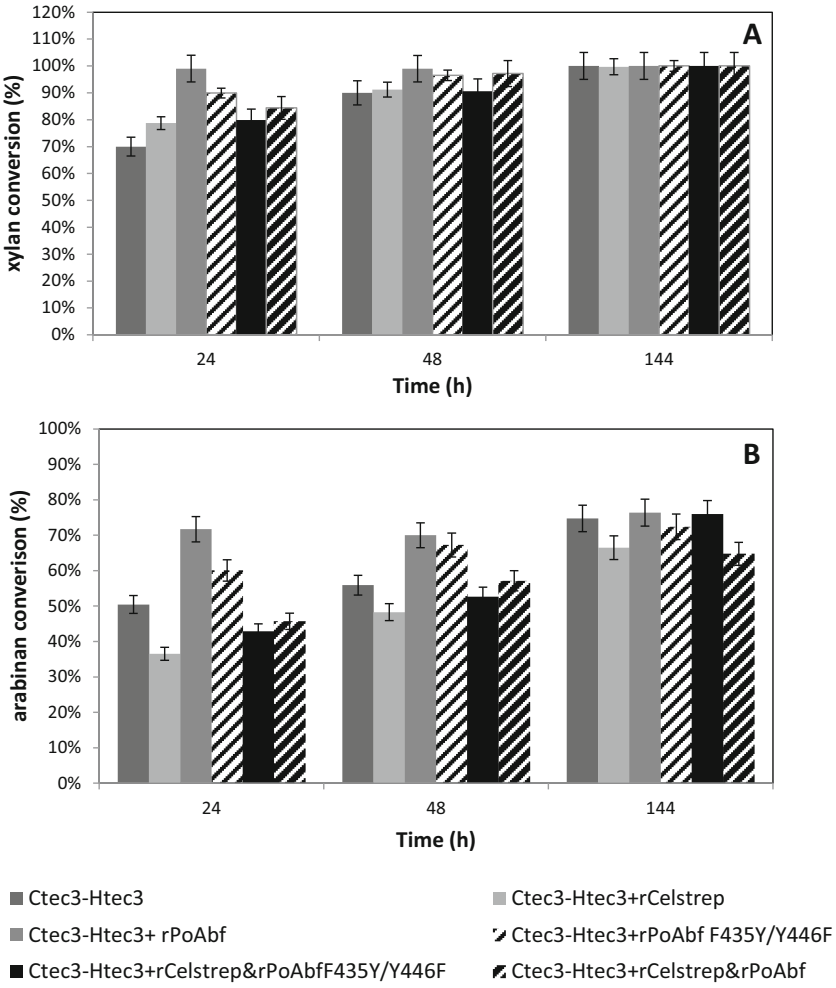


Fig. 4 Xylan (**a**) and arabinan (**b**) conversion of AFEX-treated corn stover achieved using commercial enzyme preparation Novozymes Cellic®, 60 % Ctec3 and 40 % Htec3, with or without addition of rPoabf, its variant F435Y/Y446F, rCelstrep or a combination of these

Cellic®, Ctec3 and Htec3, and rPoAbf, its evolved variant rPoAbf F435Y/Y446F and rCelStrep, no significant increase in sugar conversion was observed (Fig. 3a–c). However, it is worthy to note that supplementing rPoAbf F435Y/Y446F, the percentage of xylan and arabinan conversion increased, reaching the highest conversion (80 % for both sugars) after 6 days of hydrolysis. Comparing the data of conversion of *A. donax* pretreated by aqueous ammonia soaking (Marcolongo et al. 2014), similar xylan conversion was achieved by using the commercial mix in combination with rPoAbf and its evolved mutant after 72 h of hydrolysis. Even if the xylan hydrolysis described by Marcolongo et al. (2014) appeared to be faster than that reported in this work, it is worth to note that the amount of arabinofuranosidase enzymes in this manuscript (1.5 U/g expressed as units of purified enzyme, assayed against the substrate pNP-a-L-arabinofuranoside as previously described in Amore et al. 2012a, per gram of pretreated biomass) is lower than that (27.2 U/g) used in Marcolongo et al. (2014).

Also, in the case of corn stover, positive effects were observed in xylan and arabinan conversion when the enzymes tested in this work were added to the commercial enzyme cocktail Novozymes Cellic®, Ctec3 and Htec3. The 99 % xylan conversion was achieved in 24 h when rPoAbf wild type was supplemented to the commercial enzymes. The same xylan conversion was achieved using commercial preparation and commercial preparation in combination with rPoAbf F435Y/Y446F or rCelstrep or both enzymes after 6 days of hydrolysis. Interestingly, these data highlight that by adding rPoAbf wild type, the hydrolysis of corn stover seems to be faster than that with other preparations. These results could be due to the diverse degree of synergism among the enzymes due to different substrate proprieties as previously reported (Andersen et al. 2008; Van Dyk and Pletschke 2012).

All these data confirm the importance of using accessory enzymes like arabinofuranosidases to achieve high xylan conversion yield since they increase accessibility of xylan acting on side chain of hemicellulose and demonstrate the effectiveness of the arabinofuranosidase from the fungus *P. ostreatus* rPoAbf and, even more, of its evolved mutant rPoAbf F435Y/Y446F. It is worth to note that the use of the purified enzyme cocktail containing rPoAbf F435Y/Y446F allowed achieving higher or similar glucan, xylan and arabinan conversion after 72 h of hydrolysis (62, 63 and 80 %) than those

achieved after the same time of hydrolysis with commercial enzymes Novozymes Cellic®, Ctec3 and Htec3 (44, 66 and 55 %).

Acknowledgments This work was supported by a grant from the Ministero dell'Università e della Ricerca Scientifica—Industrial Research Project "Integrated agro-industrial chains with high energy efficiency for the development of eco-compatible processes of energy and biochemicals production from renewable sources and for the land valorization (EnerbioChem)" PON01_01966, funded in the frame of Operative National Programme Research and Competitiveness 2007–2013 D. D. Prot. n. 01/Ric. 18.1.2010. We thank Lucigen enzyme company for supplying the research enzymes and Novozyme for supplying Cte3 and Htec3 enzymes for this work. The authors also thank Dr. Valeria Ventorino of the Department of Agriculture, University of Naples "Federico II", for the support in the statistical treatment of data.

Compliance with ethical standards

Conflict of interest The authors declare that they have not competing interests.

References

Amore A, Amoresano A, Birolo L, Henrissat B, Leo G, Palmese A, Faraco V (2012a) A family GH51 α-L-arabinofuranosidase from *Pleurotus ostreatus*: identification, recombinant expression and characterization. Appl Microbiol Biotechnol 94(4):995–1006

Amore A, Pepe O, Ventorino V, Birolo L, Giangrande C, Faraco V (2012b) Cloning and recombinant expression of a cellulase from the cellulolytic strain *Streptomyces* sp. G12 isolated from compost. Microb Cell Fact 11:164

Andersen N, Johansen KS, Michelsen M, Stenby EH, Krogh KB, Olsson L (2008) Hydrolysis of cellulose using mono-component enzymes shows synergy during hydrolysis of phosphoric acid swollen cellulose (PASC), but competition on Avicel. Enzyme Microb Technol 42:362–370

Balan V, Da Costa SL, Chundawat SPS, Marshall D, Sharma LN, Chambliss CK, Dale BE (2009) Enzymatic digestibility and pretreatment degradation products of AFEX-treated hardwoods (*Populus nigra*). Biotechnol Prog 25:365–375

Bals B, Rogers C, Jin M, Balan V, Dale BE (2010) Evaluation of ammonia fibre expansion (AFEX) pretreatment for enzymatic hydrolysis of switchgrass harvested in different seasons and locations. Biotechnol Biofuels 3:1

Boisset C, Petrequin C, Chanzy H, Henrissat B, Schulein M (2001) Optimized mixtures of recombinant *Humicola insolens* cellulases for the biodegradation of crystalline cellulose. Biotechnol Bioeng 72(3):339–45

Caparrós S, Ariza J, Lopez F, Diaz MJ (2007) Optimizing cellulosic paper obtained from *Arundo donax* L. under hydrothermal treatment. J Ind Eng Chem 13(3):465–473

De Bari I, Liuzzi F, Villone A, Braccio G (2013) Hydrolysis of concentrated suspensions of steam pretreated *Arundo donax*. Appl Energy 102:179–189

Diodato N, Fagnano M, Alberico I (2009) CliFEM–climate forcing and erosion response modelling at long-term Sele River research basin (southern Italy). Nat Hazards Earth Syst Sci 9:1693–1702

Diodato N, Fagnano M, Alberico I (2011) Geospatial–and–visual modeling for exploring sediment source areas across the Sele river landscape, Italy. Ital J Agron 6:85–92

Fagnano M, Diodato N, Alberico I, Fiorentino N (2012) An overview of soil erosion modeling compatible with RUSLE approach. Rend Fis Acc Lincei 23:69–80

Fagnano M, Impagliazzo A, Mori M, Fiorentino N (2015) Agronomic and environmental impacts of giant reed (*Arundo donax* L.): results from a long-term field experiment in hilly areas subject to soil erosion. Bioenergy Res 8:415–422

Fiorentino N, Fagnano M, Adamo P, Impagliazzo A, Mori M, Pepe O, Pepe O, Ventorino V, Zoina A (2013) Assisted phytoextraction of heavy metals: compost and *Trichoderma* effects on giant reed (*Arundo donax* L.) uptake and soil N-cycle microflora. Ital J Agron 8:244–254

Fiorentino N, Impagliazzo A, Ventorino V, Pepe O, Piccolo A, Fagnano M (2010) Biomass accumulation and heavy metal uptake of giant reed on polluted soil in southern Italy. J Biotechnol 150(S1):261

Forte A, Zucaro A, Fagnano M, Bastianoni S, Basosi R, Fierro A (2015) LCA of *Arundo donax* L. lignocellulosic feedstock production under Mediterranean conditions: a three-year study. Biomass Bioenergy 73:32–47

Gao D, Chundawat SPS, Krishnan C, Balan V, Dale BE (2010a) Mixture optimization of six core glycosyl hydrolases for maximizing saccharification of ammonia fiber expansion (AFEX) pretreated corn stover. Bioresour Technol 101:2770–2781

Gao D, Chundawat SPS, Liu T, Hermanson S, Gowda K, Brumm P, Dale BE, Balan V (2010b) Strategy for identification of novel fungal and bacterial glycosyl hydrolase hybrid mixtures that can efficiently saccharify pretreated lignocellulosic biomass. Bioenergy Res 3:67–81

Gao D, Uppugundla N, Chundawat SPS, Yu X, Hermanson S, Gowda K, Brumm P, Mea D, Balan V, Dale BE (2011) Hemicellulases and auxiliary enzymes for improved conversion of lignocellulosic biomass to monosaccharides. Biotechnol Biofuels 4(1):5

Giacobbe S, Vincent F, Faraco V (2014) Development of an improved variant of GH51 α-l-arabinofuranosidase from *Pleurotus ostreatus* by directed evolution. New Biotechnol 31(3):230–236

Hames B, Ruiz R, Scarlata C, Sluiter A, Sluiter J, Templeton D (2008) Preparation of samples for compositional analysis. National Renewable Energy Laboratory, Golden, CO, USA: NREL Laboratory Analytical

Harun S, Balan V, Takriff MS, Hassan O, Jahim J, Dale BE (2013) Performance of AFEX pretreated rice straw as source of fermentable sugars: the influence of particle size. Biotechnol Biofuels 6:40

Hermann BG, Blok K, Patel MK (2007) Producing bio-based bulk chemicals using industrial biotechnology saves energy and combats climate change. Environ Sci Technol 41(22):7915–21

Kajaste R (2014) Chemicals from biomass e managing greenhouse gas emissions in biorefinery production chains e a review. J Clean Prod. doi:10.1016/j.jclepro.2014.03.070

Kobayashi H, Fukuoka A (2013) Synthesis and utilization of sugar compounds derived from lignocellulosic biomass. Green Chem 15(7):1740–1763

Kumar P, Barrett DM, Delwiche MJ, Stroeve P (2009) Methods for pretreatment of lignocellulosic biomass for efficient hydrolysis and biofuel production. Ind Eng Chem Res 48:3713–3729

Li C, Cheng G, Balan V, Kent MS, Ong M, Chundawat SPS, daCosta SL, Melnichenko YB, Dale BE, Simmons BA, Singh S (2011) Influence of physico-chemical changes on enzymatic digestibility of ionic

liquid and AFEX pretreated corn stover. Bioresour Technol 102(13): 6928–6936

Marcolongo L, Ionata E, La F, Amore A, Giacobbe S, Pepe O, Faraco V (2014) The effect of *Pleurotus ostreatus* arabinofuranosidase and its evolved variant in lignocellulosic biomasses conversion. Fungal Genet Biol 72:162–167

Morrison DG (2014) Lignocellulosic waste degradation using enzyme synergy with commercially available enzymes and *Clostridium cellulovorans* xylanaseA and mannanase A. Thesis Masters M.Sc. http://hdl.handle.net/10962/d1013292

Murnen HK, Balan V, Chundawat SPS, Bals B, Sousa LD, Dale BE (2007) Optimization of ammonia fiber expansion (AFEX) pretreatment and enzymatic hydrolysis of *Miscanthus x giganteus* to fermentable sugars. Biotechnol Prog 23:846–850

Polizeli MLTM, Rizzatti ACS, Monti R, Terenzi HF, Jorge JA, Amorim DS (2005) Xylanases from fungi: properties and industrial applications. Appl Microbiol Biotechnol 67(5):577–591

Scheidel A, Sorman AH (2012) Energy transitions and the global land rush: ultimate drivers and persistent consequences. Glob Environ Chang 22:588–595

Scordia D, Cosentino SL, Lee JW, Jeffries TW (2012) Bioconversion of giant reed (*Arundo donax* L.) hemicellulose hydrolysate to ethanol by *Scheffersomyces stipitis* CBS6054. Biomass Bioenergy 39: 296–305

Sluiter A, Hames B, Ruiz R, Scarlata C, Sluiter J, Templeton D, Crocker D (2005a) Determination of structural carbohydrates and lignin in biomass. National Renewable Energy Laboratory, Golden, CO, USA: NREL Laboratory Analytical

Sluiter A, Hames B, Ruiz R, Scarlata C, Sluiter J, Templeton D (2005b) Determination of ash in biomass. National Renewable Energy Laboratory, Golden, CO, USA: NREL Laboratory Analytical Procedures

Sluiter A, Hames B, Hyman D, Payne D, Ruiz R, Scarlata C, Sluiter J, Templeton D, Wolfe J (2008) Determination of total solids in biomass and total dissolved solids in liquid process samples. National Renewable Energy Laboratory, Golden, CO, USA: NREL Laboratory Analytical

Taherzadeh MJ, Keikhosro K (2007) Enzyme-based hydrolysis processes for ethanol from lignocellulosic materials: a review. Bioresources 2: 707–738

Uppugundla N, daCosta SL, Chundawat SPS, Yu X, Simmons B, Singh S, Gao X, Kumar R, Wyman CE, Dale BE, Balan V (2014) A comparative study of ethanol production using dilute acid, ionic liquid and AFEX™ pretreated corn stover. Biotechnol Biofuels 13:7–72

Van Den Brink J, Van Muiswinkel GCJ, Theelen B, Hinz SW, De Vries RP (2013) Efficient plant biomass degradation by thermophilic fungus *Myceliophthora heterothallica*. Appl Environ Microbiol 79(4): 1316–1324

Van Dyk JS, Pletschke BI (2012) A review of lignocellulose bioconversion using enzymatic hydrolysis and synergistic cooperation between enzymes—factors affecting enzymes, conversion and synergy. Biotechnol Adv 30:1458–1480

Wettstein SG, Alonso DM, Gu EI, Dumesic JA (2012) A roadmap for conversion of lignocellulosic biomass to chemicals and fuels. Curr Opin Chem Eng 1:218–224

Woodward J (1991) Synergism in cellulase systems. Biores Technol 36: 67–75

2.2 – Upgrading of Municipal Solid Waste. The case of Newspaper Waste

The world Municipal Solid Waste (MSW) volume is expected to reach 2.2 billion tonnes by 2025, due to the increase of urban residents (to 4.3 billion), generating about 1.42 kg/capita/day of MSW. Improperly managed waste is one of the main causes of a serious health hazard and nowadays environmental concerns are increasing the attention on processes developed for adding value to wastes. About 50% of MSW (before recycling) by weight is paper and it is the second most abundant lignocellulose containing fraction after humid fraction. The newspaper alone constitutes up to 14% of MSW (Subhedar et al., 2015), representing one of the most attractive feedstock in biorefinering industry, since it contains up to 70% polysaccharides (Wang et al., 2012). However, most types of waste paper can contain not only lignin, but also additional barrier components such as waxes, fats, rubbers, synthetic polymers and resins, and other hydrophobic substances that prevent the enzymatic cleavage of the paper materials. In order to remove the barrier components and improve the enzymatic digestibility, this fraction of MSW should be subjected to an appropriate pretreatment method.

The following paper **“Saccharification of newspaper waste after ammonia fiber expansion or extractive ammonia” (PAPER II)** describes the upgrading evaluation of Newspaper Waste as feedstock for valuable bioproducts and biochemicals production via fermentation. This objective was achived by testing two different physico-chemical pretreatment methods – Ammonia Fiber Expansion (AFEX) and Extractive Ammonia (EA). This research was supported by a Marie Curie International Research Staff Exchange Scheme Fellowship within the 7th European Community Framework Programme: ‘Improvement of technologies and tools, e.g., biosystems and biocatalysts, for waste conversion to develop an assortment of high added value eco-friendly and cost-effective bio-products’ BIOASSORT (grant number 318931).

Montella *et al. AMB Expr (2016) 6:18*
DOI 10.1186/s13568-016-0189-9

AMB Express
a SpringerOpen Journal

ORIGINAL ARTICLE

Open Access

Saccharification of newspaper waste after ammonia fiber expansion or extractive ammonia

Salvatore Montella[1], Venkatesh Balan[2], Leonardo da Costa Sousa[2], Christa Gunawan[2], Simona Giacobbe[1], Olimpia Pepe[3] and Vincenza Faraco[1*]

Abstract

The lignocellulosic fractions of municipal solid waste (MSW) can be used as renewable resources due to the widespread availability, predictable and low pricing and suitability for most conversion technologies. In particular, after the typical paper recycling loop, the newspaper waste (NW) could be further valorized as feedstock in biorefinering industry since it still contains up to 70 % polysaccharides. In this study, two different physicochemical methods—ammonia fiber expansion (AFEX) and extractive ammonia (EA) were tested for the pretraetment of NW. Furthermore, based on the previously demonstrated ability of the recombinant enzymes endocellulase rCelStrep, α-L-arabinofuranosidase rPoAbf and its evolved variant rPoAbf F435Y/Y446F to improve the saccharification of different lignocellulosic pretreated biomasses (such as corn stover and *Arundo donax*), in this study these enzymes were tested for the hydrolysis of pretreated NW, with the aim of valorizing the lignocellulosic fractions of the MSW. In particular, a mixture of purified enzymes containing cellulases, xylanases and accessory hemicellulases, was chosen as reference mix and rCelStrep and rPoAbf or its variant were replaced to EGI and Larb. The results showed that these enzymatic mixes are not suitable for the hydrolysis of NW after AFEX or EA pretreatment. On the other hand, when the enzymes rCelStrep, rPoAbf and rPoAbf F435Y/Y446F were tested for their effect in hydrolysis of pretreated NW by addition to a commercial enzyme mixture, it was shown that the total polysaccharides conversion yield reached 37.32 % for AFEX pretreated NW by adding rPoAbf to the mix whilst the maximum sugars conversion yield for EA pretreated NW was achieved 40.80 % by adding rCelStrep. The maximum glucan conversion yield obtained (45.61 % for EA pretreated NW by adding rCelStrep to the commercial mix) is higher than or comparable to those reported in recent manuscripts adopting hydrolysis conditions similar to those used in this study.

Keywords: AFEX pretreatment, Arabinofuranosidase, Biorefining, Cellulase, EA pretreatment, Hemicellulase, Municipal solid waste, Newspaper waste

Introduction

The world municipal solid waste (MSW) volume is expected to reach 2.2 billion tonnes by 2025, due to the increase of urban residents to 4.3 billion. Improperly managed urban solid waste is one of the main causes of environmental pollution and a serious health hazard, due to contamination of groundwater and surface water by leachate, as well as air pollution by their burning. Monitoring of pollution from different waste management options is crucial and environmental concerns are increasing the focus on the reuse and the recycling of MSW and on processes for adding value to wastes. In particular, the lignocellulosic fractions of MSW can be a source of a wide range of high added value products by biotechnological "tailor made" processes (Liguori et al. 2013) for the biorefineries development (Menon and Rao 2012; Fava et al. 2015; Esposito and Antonietti 2015). The future of biorefining industry depends mostly on the availability of cheap, sustainable and abundant biomasses

*Correspondence: vfaraco@unina.it
[1] Department of Chemical Sciences, University of Naples "Federico II", Complesso Universitario Monte S. Angelo, Via Cintia, 4, 80126 Naples, Italy
Full list of author information is available at the end of the article

as feedstock. The lignocellulosic fractions from MSW are more suitable to achieve this objective than dedicated agricultural crops (avoiding the displacement of food crops and minimizing the conflict food versus fuel) or forest resources (avoiding the issues concerning a massive deforestation). Moreover the change of land use and the removal of forest trees in large areas can have a negative impact on the ecosystem. The lignocellulosic fractions from MSW show further advantages such as widespread availability, predictable and low pricing, and suitability for most conversion technologies.

Among the lignocellulosic fractions of MSW, paper waste represents approximately 30 % of MSW and it is the second most abundant lignocellulose fraction after humid fraction (U.S. Environmental Protection Agency—EPA 2015). The major method of paper waste management is recycling: according to EPA, the paper waste can typically be recycled from 5 to 7 times, before papermaking fibers become too short and weak to hold together; each recycling requires de-inking with chemicals processing and by adding virgin wood fibers. The newspaper waste (NW) alone constitutes up to 14 % of MSW (Subhedar et al. 2015), and it can be recycled fewer times that office paper due to the fact that it is usually made by shorter fibers. After the typical paper recycling loop, the NW can be further valorized as feedstock in biorefinering industry since it still contains up to 70 % polysaccharides which can be hydrolysed to fermentable sugars (Wang et al. 2012a).

For an effective saccharification, the NW, as all the lignocellulose biomasses, requires a physical, chemical or enzymatic pretreatment to break down the recalcitrant lignin and increase the polysaccharides accessibility for the following hydrolysis. The latter one can be performed by enzymatic method, that is more eco-friendly than chemical conversion, but it is not yet economically competitive, mainly due the high cost of the needed enzymes. Development of enzymes with improved performances by enhancing their stability and specific activity is therefore pursued. Unlike dedicated energy crops and agro-industrial waste, the NW saccharification yield is generally low probably due to high lignin content, dense structure and the additional physical barrier constituted by toner's ink and inorganic coating linked to the lignin (Chu and Feng 2013; Kim et al. 2006; Kuhad et al. 2010). In order to overcome this bottleneck, the evaluation of the most efficient pretreatment method and of a suitable tailor-made enzymatic mixture are the crucial steps.

This study is aimed at evaluating the feasibility of NW as feedstock for the fermentable sugars production by testing two different physicochemical pretreatment methods, ammonia fiber expansion (AFEX) and extractive ammonia fiber expansion (EA).

We have previously shown that tailor-made enzymatic cocktails including the endocellulase rCelStrep from *Streptomyces* sp. G12 recombinantly expressed in *Escherichia coli* (Amore et al. 2012b), the α-L-arabinofuranosidase rPoAbf from the fungus *Pleurotus ostreatus* recombinantly expressed in *Pichia pastoris* (Amore et al. 2012a) and/or its evolved variant rPoAbf F435Y/Y446F (Giacobbe et al. 2014) are suitable to improve the hydrolysis yields of AFEX pretreated corn stover and *Arundo donax* (Giacobbe et al. 2015). Based on these results, these enzymes were chosen in this study to investigate their ability to improve the saccharification of AFEX- and EA-treated NW, with the aim of valorizing this lignocellulosic MSW fraction.

Materials and methods

Feedstock

NW was used as feedstock for the saccharification experiments. The NW was collected in the recycle station of the Office for International Student and Scholar (OISS) at Michigan State University (MSU), East Lansing, Michigan and it was mainly composed of free weekly newspaper distributed in Lansing area. After shredding into pieces of 1 cm in wide and 20 cm in length, the NW was milled with a 2 mm diameter sieve and stored under dry conditions at room temperature until use.

Compositional analysis

The composition analyses of NW were performed by acid hydrolysis according to the laboratory analytical procedures (LAPs) developed by the National Renewable Energy Laboratory (NREL) (Sluiter et al. 2010; Templeton et al. 2010):

- "Preparation of samples for compositional analysis" (Hames et al. 2008),
- "Determination of structural carbohydrates and lignin in biomass" (Sluiter et al. 2008c),
- "Determination of total solids in biomass and total dissolved solids in liquid process samples" (Sluiter et al. 2008a),
- "Determination of ash in biomass" (Sluiter et al. 2008b).

A moisture analyser was used to evaluate the moisture content. The acid insoluble lignin (Klason lignin) was detected by weighting the dried residue after total removal of the sugars. Monomeric sugars were quantified using a Biorad Aminex HPX-87H high-performance liquid chromatography (HPLC) column using 5 mM sulphuric acid as mobile phase.

AFEX pretreatment

NW was subjected to AFEX pretreatment by varying reaction temperature (65–75 °C), moisture (10.7–25 % on dry weight basis), ammonia to biomass ratio (2.0:1 and 2.8:1) at fixed residence time of 15 min. AFEX was done in a high pressure stainless steel Parr reactor. Biomass was sprayed with water to reach appropriate moisture content after taking into consideration the moisture content of original biomass. Then the reactor was closed and vacuum applied to remove residual air in the reactor. The required amount of liquid ammonia was loaded using a ammonia delivery pump into a reactor pre-heated by external heating mantle. The biomass was mixed during AFEX pretreatment process for 15 min. As the temperature of reactor was increased, the pressure in the vessel increased (between 200 and 400 psi) relying on the ammonia to biomass loading. After the completion of pretreatment process, the pressure was released from the vessel and ammonia was vented in the hood. Subsequently, the biomass was transferred to a tray and dried in the hood overnight to remove residual ammonia. The dry pretreated biomass was stored in a sealed polythene bag at 4 °C until use.

Extractive ammonia (EA) pretreatment

NW was subjected to EA pretreatment (reaction temperature: 120 °C; ammonia to biomass ratio: 3:1 kg/kg dry biomass; moisture content: 10 % on dry weight basis; fixed residence time: 15 min). EA was done in a high pressure stainless steel Parr reactor. Biomass was first sprayed with water to reach appropriate moisture content after taking into consideration the moisture content of original biomass. The required amount of liquid ammonia was loaded using a ammonia delivery pump into a reactor pre-heated by external heating mantle. The biomass was mixed during EA pretreatment process for 15 min. As the temperature of reactor was increased, the pressure in the vessel raised (between 600 and 800 psi) depending on the ammonia to biomass loading. After the completion of pretreatment process, extractives (mostly lignin) generated during the process along with liquid ammonia was collected via a sintered frit into another high pressure reactor. The collection pressure was released from the vessel and ammonia was vented in the hood to recover the extractives. The pretreated biomass was transferred to a tray and dried in the hood overnight to remove residual ammonia present in the biomass. Then dry EA treated biomass was stored in a sealed polythene bag in a refrigerator until further use.

Commercial and purified enzymes and their sources

The commercially available complex of cellulases and hemicellulases Cellic® CTec3 and the enzyme solution Cellic® HTec3 from Novozymes (Denmark) were used in this study. Moreover, mixes of the following purified enzymes were tested. The four core fungal cellulases were cellobiohydrolase I (CBH I; glycoside hydrolase—GH—family 7A), cellobiohydrolase II (CBH II; GH family 6A), endoglucanase I (EG I; GH family 7B) and β-glucosidase (βG; GH family 3). CBH I, CBH II and EG I were purified from Spezyme CP (Danisco US Inc., Genencor Division, Rochester, NY) using four different chromatography methods: size exclusion, anion and cation exchange, hydrophobic interaction and affinity (Gao et al. 2010a, b); βG was purified from Novo 188 (Novozyme, Davis, CA, USA), by using anion and cation exchange chromatography (Gao et al. 2010a, b). The bacterial hemicellulases added to the core cellulases were xylanases (LX3, GH family 10; LX4, GH family 11), β-xylosidase (LβX; GH family 52) and α-arabinofuranosidase (LArb, GH family 51). LX3 and LX4 from *Clostridium thermocellum*, LβX from *Geobacillus stearothermophilus* and LArb from *Geobacillus* sp. G11MC16 were recombinantly expressed in *E. coli* BL21 (DE3) and purified using HIS-select nickel affinity chromatography (Gao et al. 2010a, b, 2011). The endoglucanase rCelStrep (GH family 12) was from *Streptomyces* sp. G12 and recombinantly expressed in *E. coli* (Amore et al. 2012b); the α-L-arabinofuranosidases rPoAbf wild type (GH family 51) from *P. ostreatus* and its mutant rPoAbf F435Y/Y446F were recombinantly expressed in *P. pastoris* (Amore et al. 2012a). These enzymes were purified by ammonium sulphate precipitation followed by hydrophobic interaction chromatography (Amore et al. 2012a, b).

Enzymatic hydrolysis

The enzymatic hydrolyses were carried in 5 mL vials. An amount of pretreated biomass was added as to load 1 % (w/w) glucan in 2 mL total volume with the desired enzymes. The buffer solution was 50 mM citrate, pH 4.8. Microbial and fungal contaminations were prevented by adding sodium azide 0.5 mM. The hydrolysis parameters were: 50 °C, 250 rpm, 72 h. Sampling was collected every 24 h to evaluate carbohydrates hydrolysis.

Fifteen milligram per gram of glucan of the commercial enzymatic preparation Novozymes Cellic® (60 % CTec3 and 40 % HTec3) was used for hydrolysis experiment to select the best pretreatment conditions.

The following enzymes were used for the hydrolysis of AFEX and EA pretreated NW. The MIX A was prepared including: CBH I, CBH II and EG I (3.32 mg/g glucan each), βG (2 mg/g glucan), LX3 and LX4 (1.66 mg/g glucan each), LβX and LArb (0.6 mg/g glucan each).

In the mix B, the EG I was replaced by an equal amount of endoglucanase rCelStrep; in the mix C the LArb was replaced by an equal amount of the α-L-arabinofuranosidase rPoAbf; in the mix D the LArb

was replaced by the same amount of the mutant rPoAbf F435Y/Y446F.

Moreover, rCelStrep, rPoAbf or rPoAbf F435Y/Y446F were added alternatively or in combination to the MIX 1 containing 15 mg/g glucan of commercial preparation mix (60 % Cellic® CTec3 and 40 % Cellic® HTec3 from Novozymes). In particular the following enzymes were added: in the MIX 2, the endoglucanase rCelStrep (3.32 mg/g glucan); in the MIX 3 the α-L-arabinofuranosidase rPoAbf (0.6 mg/g glucan); in the MIX 4 the mutant rPoAbf F435Y/Y446F (0.6 mg/g glucan); in the MIX 5 both the endoglucanase rCelStrep (3.32 mg/g glucan) and the α-L-arabinofuranosidase rPoAbf (0.6 mg/g glucan); in the MIX 6 both the endoglucanase rCelStrep (3.32 mg/g glucan) and the mutant α-L-arabinofuranosidase rPoAbf F435Y/Y446F (0.6 mg/g glucan).

Sugar analysis

Monomeric sugars concentration were determined by high performance liquid chromatography (HPLC). All experiments were performed in triplicate.

About 200 µl hydrolysate were collected in a centrifuge tube, heated to 100 °C for 10 min (to inactivate the enzymes), then spun down at 8000 rpm for 10 min and the supernatant was stored in a HPLC vial at −20 °C until further use. Monomeric sugars concentration in the hydrolysate was determined by HPLC using Biorad Aminex HPX-87P. Shimadzu HPLC Prominence system (Columbia, MD, USA) with a refractive index detector (RID) were used for analyzing the sugars. Water was used as the mobile phase at a fixed flow rate of 0.6 ml/min, with isocratic elution. The column temperature was maintained at 60 °C and the HPLC sample injection volume was 20 µl. Standard curves were generated using different concentrations of mixed sugars. A guard column with similar packing was used throughout the chromatography experiments.

The sugars conversion is calculated according to the following equations:

$$\text{Glucan conversion } (\%) = \frac{\text{Glucose concentration in hydrolysate } (g/L)}{\text{Glucan concentration in loaded biomass } (g/L)} \times \frac{1}{1.11}$$

$$\text{Xylan conversion } (\%) = \frac{\text{Xylose concentration in hydrolysate } (g/L)}{\text{Xylan concentration in loaded biomass } (g/L)} \times \frac{1}{1.136}$$

$$\text{Arabinan conversion } (\%) = \frac{\text{Arabinose concentration in hydrolysate } (g/L)}{\text{Arabinan concentration in loaded biomass } (g/L)} \times \frac{1}{1.136}$$

$$\text{Mannan conversion } (\%) = \frac{\text{Mannose concentration in hydrolysate } (g/L)}{\text{Mannan concentration in loaded biomass } (g/L)} \times \frac{1}{1.11}$$

where 1.11 is the ratio of $MW_{glucose\ (or\ mannose)}$ (180.16 g/mol) to $MW_{glucan\ (or\ mannan)}$ (162.14 g/mol) and 1.11 is the ratio of $MW_{xylose\ (or\ arabinose)}$ (150.13 g/mol) to $MW_{xylan\ (or\ arabinan)}$ (132.13 g/mol).

Results

Composition analysis and pretreatment of newspaper waste

Composition analysis of untreated NW was carried out and the results reported in Table 1 revealed a glucan content of ~44 %, a hemicellulose content of ~15 % and a Klason lignin content of ~25 %.

The macromolecular composition of AFEX pretreated NW is assumed to have the same composition of untreated NW (Table 1), due to the fact that, as previously reported, the AFEX pretreatement preserves the macrostructure of the lignocellulosic biomasses, reducing the degree of polymerization of (hemi)cellulose minimizing the degradation of the original carbohydrates (Holtzapple et al. 1991; Kumar et al. 2009). In order to select the best AFEX conditions to be used during the further saccharification experiments, three different sets of conditions were tested on NW (Table 2), varying ammonia loading, temperature and moisture content. These conditions were chosen based on the previous published screening of AFEX pretreatments on NW (Holtzapple et al. 1991, 1992).

The AFEX pretreated NW was subjected to enzymatic hydrolysis using the mix composed of 60 % Cellic® CTec3 and 40 % Cellic® HTec3 from Novozymes. As shown in Fig. 1, the maximum glucose and xylose yields were achieved after 72 h of hydrolysis with the AFEX pretreatment condition N1 (2.8 kg ammonia/kg dry biomass, 65 °C, and 10.7 % of moisture content) and N2 (2.8 kg ammonia/kg dry biomass, 75 °C, and 25 % of moisture content) and the maximum total sugars conversion was respectively 29.58 ± 0.35 and 29.02 ± 0.35 % (Table 2). The condition N1 was selected for the further hydrolysis experiments due to the milder reaction temperature of pretreatment (10 °C less than condition N2).

Table 1 Macromolecular composition of newspaper waste before and after EA pretreatment

	Newspaper waste	
	Untreated	**EA pretreated**
Moisture content	7.00 ± 1.00	9.75 ± 0.07
Ash	7.04 ± 0.14	7.25 ± 0.15
Structural carbohydrate		
Glucan	44.21 ± 4.02	41.36 ± 0.01
Xylan	5.11 ± 0.19	5.20 ± 0.22
Galactan	1.81 ± 0.13	1.84 ± 0.03
Arabinan	1.09 ± 0.09	1.73 ± 0.03
Mannan	9.83 ± 0.47	8.16 ± 0.05
Lignin		
Acid insoluble lignin	25.88 ± 2.48	30.03 ± 0.89
Acid soluble lignin	0.96 ± 0.01	1.05 ± 0.19
Composition closure	95.92	96.63

As an alternative pretreatment method, the EA was also tested on NW. The EA pretreatment converts native cellulose I to cellulose III, delignifying the biomass simultaneously (da Costa et al. 2015). Composition analysis of EA pretreated NW was carried out and the results showed that the NW still contains significant percentages of structural sugars after EA pretreatment (~58 %). In particular, the percentages of glucan content is ~41 %, the percentage of mannan content is ~8 % and the percentage of xylan content is ~5 %; the total lignin content is ~31 % (Table 1).

Hydrolysis of newspaper waste by using the enzymes rCelStrep, rPoAbf and its variant in comparison with the enzymes EGI or LArb

As shown in several papers, a mix of cellulases and accessory hemicellulases is crucial to enhance hydrolysis of pretreated lignocellulosic biomasses (Gao et al. 2010a, b; Jørgensen et al. 2007). The MIX A, previously optimized for corn stover saccharification after AFEX pretreatment (Gao et al. 2011), was chosen as reference mix. This mixture contains the cellobiohydrolases CBHI and CBHII, the endo-glucanase EGI, the beta-glucosidase βG, the xylanases LX3 and LX4, the beta-xylosidase LβX and the α-L-arabinofuranosidase LArb. In order to evaluate the hydrolysis yield by using the enzymes rCelStrep, rPoAbf and rPoAbf F435Y/Y446F, these enzymes were replaced to EGI (MIX B) and to LArb (MIX C and MIX D), respectively.

The highest glucan, xylan, mannose and arabinose conversions were reached after 72 h of hydrolysis for

Table 2 AFEX conditions tested on newspaper waste

AFEX conditions	Ammonia loading (kg:kg dry biomass)	Reaction temperature (°C)	Moisture content (%)	Fixed residence time (min)	Sugars conversion (after 72 h hydrolysis)	
					(%)	(g/L)
N1	2.8:1	65	10.7	15	29.58	4.62
N2	2.8:1	75	25		29.02	4.53
N3	2.0:1	75	25		26.14	4.08

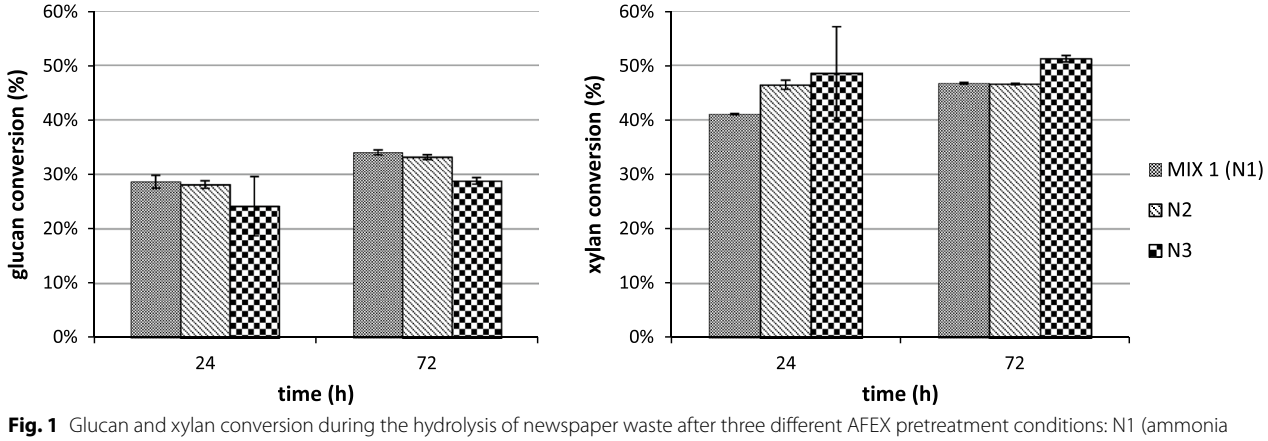


Fig. 1 Glucan and xylan conversion during the hydrolysis of newspaper waste after three different AFEX pretreatment conditions: N1 (ammonia loading 2.8:1; 65 °C; 10.7 % moisture content; 15 min), N2 (ammonia loading 2.8:1; 75 °C; 25 % moisture content; 15 min) and N3 (ammonia loading 2.0:1; 65 °C; 10.7 % moisture content; 15 min)

all the tested biomasses, pretreatment methods and enzyme mixes. In Figs. 2 and 3 the conversion data for the most abundant polysaccharides (glucan and xylan) are reported.

For AFEX pretreated NW, as shown in Fig. 2, the glucan conversion reached 7.3 % for MIX A and did not increase significantly by using the other mixes. The xylan conversion was 13.59 % for MIX A and increased up to 14.44 % only for MIX B (corresponding to the substitution of rCelStrep to EGI), decreasing for the other mixes (12.2 % for MIX C and 13.1 % for MIX D). The arabinose conversion reached 7.61 % for MIX A, increased up to 8.86 % for MIX B and decreased to 4.4 % for MIX C and to 5.3 % for MIX D. The mannose conversion was undetectable for all mixes. In conclusion, there was no substantial improvement by replacing rCelStrep, rPoAbf or its variant to the corresponding enzymatic activities in the reference mix.

As shown in Fig. 3, for EA pretreated NW, the glucan conversion reached 14.3 % for MIX A and decreased for other three mixes (up to 13.1 % for MIX B, 10.7 % for MIX C and 13.4 % for MIX D). The xylan conversion was 19.2 % for MIX A and increased only for MIX B (up to 20.3 %), decreasing for MIX C (16.9 %) and for MIX D (15.4 %). The arabinose conversion reached 12.5 % for MIX A and decreased to 11.1 % for MIX B, 9.9 % for MIX C and 8.70 % for MIX D. The mannose conversion was less than 1 % for all the mixes. In conclusion, the

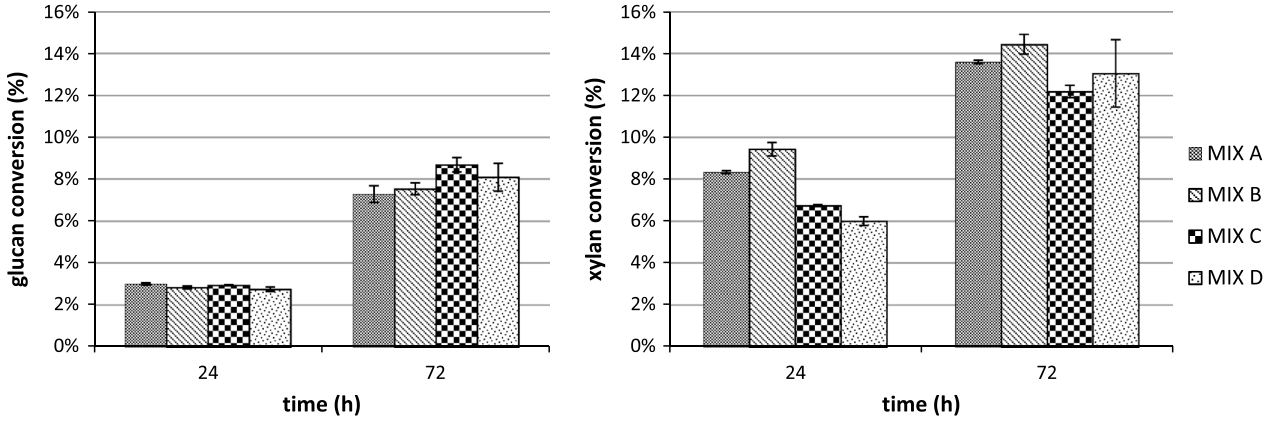


Fig. 2 Glucan and xylan conversion during the hydrolysis of AFEX pretreated newspaper waste by using MIX A: CBH I, CBH II and EG I (3.32 mg/g glucan each), βG (2 mg/g glucan), LX3 and LX4 (1.66 mg/g glucan each), LβX and LArb (0.6 mg/g glucan each); MIX B (replacing EG I with endoglucanase rCelStrep); MIX C (replacing LArb with the α-L-arabinofuranosidase rPoAbf); MIX D (replacing LArb with mutant rPoAbf F435Y/Y446F)

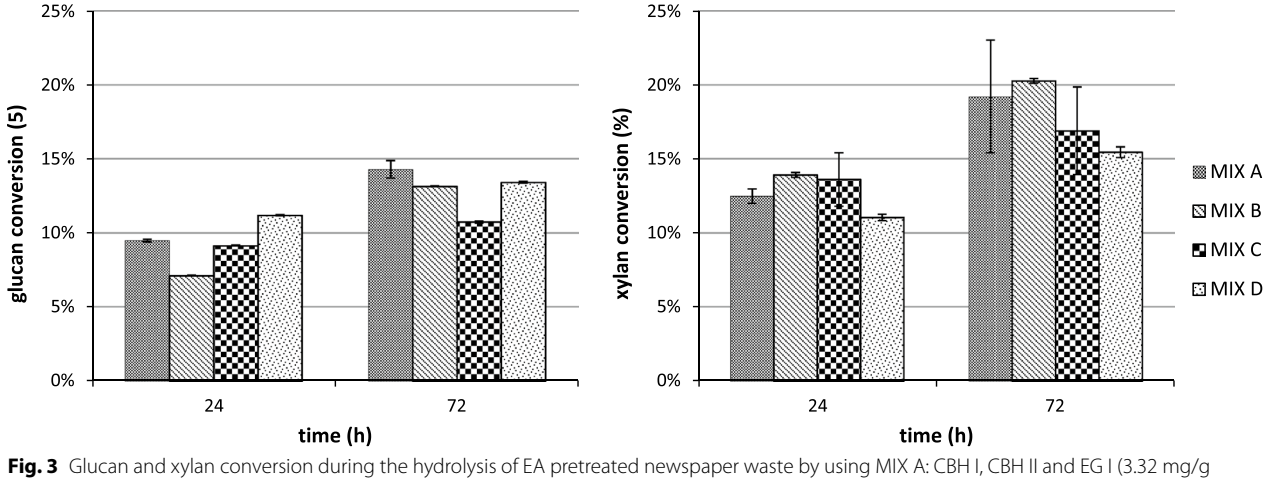


Fig. 3 Glucan and xylan conversion during the hydrolysis of EA pretreated newspaper waste by using MIX A: CBH I, CBH II and EG I (3.32 mg/g glucan each), βG (2 mg/g glucan), LX3 and LX4 (1.66 mg/g glucan each), LβX and LArb (0.6 mg/g glucan each); MIX B (replacing EG I with endoglucanase rCelStrep); MIX C (replacing LArb with the α-L-arabinofuranosidase rPoAbf); MIX D (replacing LArb with mutant rPoAbf F435Y/Y446F)

maximum polysaccharides conversion yield (12.4 %) was obtained for MIX A.

Evaluation of synergism between commercial enzyme preparation and rCelstrep and/or rPoAbf and its mutant

The enzymes rCelStrep, rPoAbf or its evolved variant rPoAbf F435Y/Y446F were added alternatively or in combination to 15 mg/g glucan of MIX 1 (60 % Cellic® CTec3 and 40 % Cellic® HTec3 from Novozymes) for the hydrolysis reactions of pretreated NW. The amount of enzyme loading was chosen based on the data obtained by Gao et al. (2011) and due to the promising results previously achieved for the hydrolysis of AFEX pretreated *Corn stover* and *A. donax* (Giacobbe et al. 2015).

The highest glucan, xylan and arabinose conversion were reached after 72 h of hydrolysis for all tested biomasses, pretreatment methods and enzyme mixes. In Figs. 4 and 5 the conversion data for the most abundant polysaccharides (glucan and xylan) are reported.

As shown in Fig. 4, for AFEX pretreated NW, the glucan conversion reached 34.0 % for MIX 1 and increased up to 39.4 % for MIX 2, 41.2 % for MIX 3, 42.7 % for MIX 4, 41.22 % for MIX 5 and 41.0 % for MIX 6. The xylan conversion was 46.7 % for MIX 1 and increased up to 53.3 % for MIX 2, 57.5 % for MIX 3, 58.7 % for MIX 4, 57.50 % for MIX 5 and 56.7 % for MIX 6. The arabinose conversion reached 80.7 % for MIX 1 and decreased for all other mixes (67.3 % for MIX 2, 71.6 % for MIX 3, 70.8 % for MIX 4, 71. 7 % for MIX 5 and 70.0 % for MIX 6). The mannose conversion was undetectable for MIX 1 and reached 4.5 % for MIX 2, 6.1 % for MIX 3, 4.7 % for mix 4, 4.6 % for MIX 5 and 4.5 % for MIX 6. In conclusion, the maximum polysaccharides conversion (37.32 %) was obtained by adding rPoAbf F435Y/Y446F (0.6 mg/g glucan) to MIX 1.

As shown in Fig. 5, for the EA pretreated NW, the glucan conversion reached 35.6 % for MIX 1 and increased up to 45.6 % for MIX 2, 40.1 % for MIX 3, 43.6 % for MIX 4, 42.657 % for MIX 5 and 38.5 % for MIX 6. The xylan conversion was 44.9 % for MIX 1 and increased up to 55.6 % for MIX 2, 51.3 % for MIX 3, 58.1 % for MIX 4, 53.7 % for MIX 5 and 49.2 % for MIX 6. The arabinose conversion reached 49.0 % for MIX 1 and increased up to 52.1 % for MIX 2, 49.2 % for MIX 3, 53.7 % for MIX 4, 51.0 % for MIX 5 and 50.2 % for MIX 6. The mannose conversion reached 17.2 % for MIX 1 and decreased for all the other mixes (13.6 % for MIX 2, 12.8 % for MIX 3, 11.4 % for MIX 4, 13.0 % for MIX 5 and 10.0 % for MIX 6. In conclusion, the maximum polysaccharides conversion yield (40.80 %) was obtained for MIX 2, by adding rCelStrep (3.32 mg/g glucan) to MIX 1.

Discussion

Composition analysis of untreated NW, assumed equal to the macromolecular composition of AFEX pretreated NW, as aforementioned, was in agreement with the values previously reported by Wang et al. (2012b, 2013), Subhedar et al. (2015), Sangkharak (2011) and Orozco et al. (2013) related to the untreated biomass. Furthermore, composition analysis of EA pretreated NW revealed that this technique, similarly to the AFEX pretreatment, does not significantly change the composition. Furthermore, the cellulose percentage after acid pretreatment, reported by Guerfali et al. (2015), and the holocellulose percentage, after alkaline pretreatment, reported by Wu et al. (2014), were comparable to data obtained

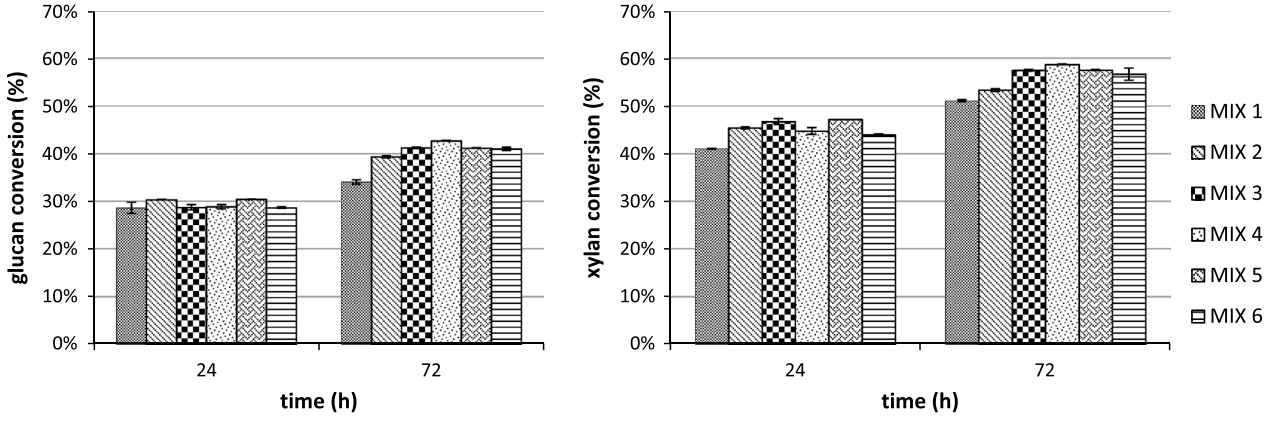


Fig. 4 Glucan and xylan conversion during the hydrolysis of AFEX pretreated newspaper waste by using MIX 1 (15 mg/g of glucan of commercial enzymatic preparation Novozymes Cellic®—60 % CTec3 and 40 % HTec3; MIX 2 (adding 3.32 mg/g glucan of endoglucanase rCelStrep to MIX 1); MIX 3 (adding 0.6 mg/g glucan of the α-L-arabinofuranosidase rPoAbf to MIX 1); MIX 4 (adding 0.6 mg/g glucan of the mutant rPoAbf F435Y/Y446F to MIX 1); MIX 5 (adding 3.32 mg/g glucan of endoglucanase rCelStrep and 0.6 mg/g glucan of the α-L-arabinofuranosidase rPoAbf to MIX 1); MIX 6 (adding 3.32 mg/g glucan of endoglucanase rCelStrep and 0.6 mg/g glucan of the mutant rPoAbf F435Y/Y446F to MIX 1)

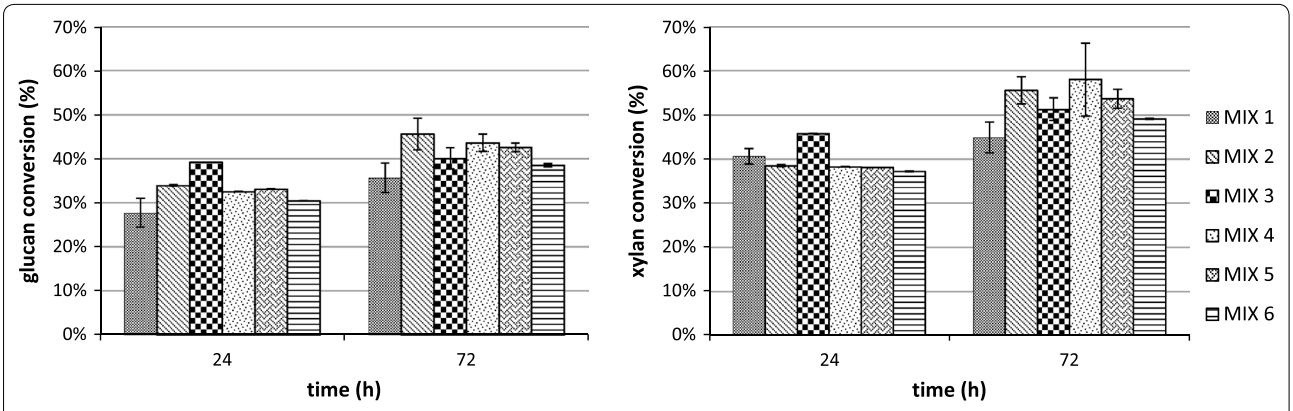


Fig. 5 Glucan and xylan conversion during the hydrolysis of EA pretreated newspaper waste by using MIX 1 (15 mg/g of glucan of commercial enzymatic preparation Novozymes Cellic®—60 % CTec3 and 40 % HTec3; MIX 2 (adding 3.32 mg/g glucan of endoglucanase rCelStrep to MIX 1); MIX 3 (adding 0.6 mg/g glucan of the α-L-arabinofuranosidase rPoAbf to MIX 1); MIX 4 (adding 0.6 mg/g glucan of the mutant rPoAbf F435Y/Y446F to MIX A'); MIX 5 (adding 3.32 mg/g glucan of endoglucanase rCelStrep and 0.6 mg/g glucan of the α-L-arabinofuranosidase rPoAbf to MIX 1); MIX 6 (adding 3.32 mg/g glucan of endoglucanase rCelStrep and 0.6 mg/g glucan of the mutant rPoAbf F435Y/Y446F to MIX 1)

after AFEX or EA method; contrariwise, the alpha cellulose and holocellulose percentage, reported by Sangkharak (2011), after pretreatment by using NaOH are both higher than those reported in this study, but the reaction conditions used for AFEX reaction are milder (65 versus 100 °C and residence time 15 min versus 4 h for AFEX and basic pretreatment respectively).

The hydrolysis of pretreated NW obtained by using mixes of purified enzymes showed that the EA method improve the saccharification in comparison with the AFEX, although the hydrolysis yields remained low in comparison to the data previously reported in literature. The maximum total polysaccharides conversion yield (for EA-pretreated-NW hydrolyzed by enzymatic MIX A) was similar to the yields reported by Ali and Khan Mohd (2011) after basic pretreatment and six days of microbial hydrolysis by *Aspergillus niger* and lower than the most recent reported results (Wua et al. 2014; Chu and Feng 2013; Wang et al. 2012b; Guerfali et al. 2015; Subhedar and Gogate 2014; Kim et al. 2006; Xin et al. 2010). Moreover, the glucan conversion yield was very low (not exceeding 7.3 %). This is the major drawback as glucose is the main carbon source for a wide range of industrial fermentation processes by using well-known microorganisms capable of using mainly hexoses (Jang et al. 2012). These results showed that the enzymatic mix previously optimized for pretreated corn stover saccharification is not suitable for the hydrolysis of NW after AFEX or EA pretreatment. Moreover, the replacement of the enzymes rCelStrep, rPoAbf and rPoAbf F435Y/Y446F respectively to EGI and LArb did not change significantly the saccharification yield. These data suggested that the purified enzymes are not probably able to overcome the physical barrier constituted by toner's ink and inorganic coating that remain linked to the lignin after pretreatment.

On the other hand, the AFEX and EA pretreatments had comparable effects on the polysaccharides conversion yield after the hydrolysis of NW by using Cellic® CTec3 and HTec3 and the addition of rCelStrep, rPoAbf or its evolved variant rPoAbf F435Y/Y446F to the commercial mix improved the sacharification process. In particular, the maximum glucan conversion yields reached for AFEX pretreated NW and EA pretreated NW were obtained by an additive effect of the α-L-arabinofuraosidase rPoAbf F435Y/Y446F and the endocellulase rCelStrep, respectively, in addition of the array of (hemi)cellulase activities present in the commercial mix. The increase of the glucan conversion yield was respectively of 25 and 28 % more than the yields obtained by hydrolysis with commercial enzymes without any addition. The best obtained sugars conversion yield was higher than or comparable to those reported in recent manuscripts adopting hydrolysis conditions similar to those used in this study. In particular, the maximum obtained glucan conversion yields were comparable to those described by Wang et al. (2012b) after enzymatic hydrolysis using 5 % (w/w) biomass and 64 mg/g glucan of commercial enzymes (~3 times more than those loaded in this study). Moreover, Wu et al. (2014) obtained sugar conversion yields comparable to this study after basic pretreatment and combining acid and enzymatic hydrolysis. Subhedar and Gogate (2014) obtained ~31 % of total sugars yield (lower than the results of this study) by using enzymatic hydrolysis and the yield increased only after an ultrasound-assisted enzymatic hydrolysis. However, the best obtained results in this study are lower than those reported by Guerfali

et al. (2015) and Kim et al. (2006) that add the action of surfactant agents to the enzymatic hydrolysis.

In conclusion, the feasibility of NW pretreatment by both AFEX and EA were demonstrated. The results showed no substantial differences between the two tested methods on hydrolysis yield. However, the AFEX was the best pretreatment technique mainly due to the mild reaction conditions. Moreover, the best sugars conversion yield obtained by adding the recombinant enzymes to the commercial mixture was higher than or comparable to those reported in previous studies adopting similar hydrolysis conditions. These promising but not optimal results suggest that the process can be optimized in order to further enhance the NW hydrolysis yield.

Authors' contributions

SM wrote the main manuscript text and prepared figures and tables. LCS designed and carried out the AFEX and EA pretreatment experiments. VB designed the hydrolysis experiments and contributed to analyze the results. SM and SG carried out the hydrolysis experiments and analyzed all the results. CG carried out the HPLC analysis. VF contributed to conceiving the study and participated in its design and coordination and drafted the manuscript. OP contributed to analyze all the results and to draft the manuscript. VF is the corresponding author. All authors read and approved the final manuscript.

Author details

[1] Department of Chemical Sciences, University of Naples "Federico II", Complesso Universitario Monte S. Angelo, Via Cintia, 4, 80126 Naples, Italy. [2] Department of Chemical Engineering and Materials Science, DOE Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, MI 48823, USA. [3] Department of Agriculture, University of Naples "Federico II", Portici, Naples, Italy.

Acknowledgements

This research was supported by a Marie Curie International Research Staff Exchange Scheme Fellowship within the 7th European Community Framework Programme: 'Improvement of technologies and tools, e.g., biosystems and biocatalysts, for waste conversion to develop an assortment of high added value eco-friendly and cost-effective bio-products' BIOASSORT (grant number 318931). We gratefully acknowledge support given to VB by the office of Biological and Environmental Research in the DOE Office of Science through the Great Lakes Bioenergy Research Centre (GLBRC) (Grant DE-FC02-07ER64494). We also thank Lucigen Enzyme Company for supplying the research enzymes and Novozyme for supplying the enzyme solutions Cellic® CTec3 and the Cellic® HTec3 for this work.

Competing interests

The authors declare that they have no competing interests.

Received: 5 December 2015 Accepted: 23 February 2016
Published online: 02 March 2016

References

Ali MN, Khan Mohd M. Production of bioethanol fuel from renewable agrobased cellulosic wastes and waste news papers. Int J Eng Sci Technol. 2011;3:884.

Amore A, Amoresano A, Birolo L, Henrissat B, Leo G, Palmese A, Faraco V. A family GH51 α-L-arabinofuranosidase from *Pleurotus ostreatus*: identification recombinant expression and characterization. Appl Microbiol Biotechnol. 2012a;94(4):995–1006.

Amore A, Pepe O, Ventorino V, Birolo L, Giangrande C, Faraco V. Cloning and recombinant expression of a cellulase from the cellulolytic strain *Streptomyces* sp.12 isolated from compost. Microb Cell Fact. 2012b;11:164–75.

Chu KH, Feng X. Enzymatic conversion of newspaper and office paper to fermentable sugars. Process Saf Environ Prot. 2013;91:123–30.

da Costa Sousa L, Jina M, Chundawata S, Bokade V, Tangd X, Azarpira A, Lu F, Avcif U, Humpula J, Uppugundlaa N, Gunawana C, Pattathilf S, Chehg A, Kotharih N, Kumarh R, Ralph J, Hahnf MG, Wymanh CE, Singh S, Simmonsi BA, Dale BE, Balan V. Next-generation ammonia pretreatment enhances biofuel production from biomass via simultaneous cellulose decrystallization and lignin extraction. Accepted: Energy and Environmental Science; 2015.

Environmental Protection Agency—EPA (2015). http://www3.epa.gov/epawaste/nonhaz/municipal/ Accessed 15 Jan 2016.

Esposito D, Antonietti M. Redefining biorefinery: the search for unconventional building blocks for materials. Chem Soc Rev. 2015;44:5821–35.

Fava F, Totaro G, Diels L, Reis M, Duarte J, Carioca OB, Poggi-Varaldo HM, Ferreira BS. Biowaste biorefinery in Europe: opportunities and research and development needs. New Biotechnol. 2015;32(1):100–8.

Gao D, Chundawat SP, Krishnan C, Balan V, Dale BE. Mixture optimization of six core glycosyl hydrolases for maximizing saccharification of ammonia fiber expansion (AFEX) pretreated corn stover. Bioresour Technol. 2010a;101(8):2770–81.

Gao D, Chundawat SP, Liu T, Hermanson S, Gowda K, Brumm P, Dale BE, Balan V. Strategy for identification of novel fungal and bacterial glycosyl hydrolase hybrid mixtures that can efficiently saccharify pretreated lignocellulosic biomass. Bioenerg Res. 2010b;3(1):67–81.

Gao D, Uppugundla N, Chundawat SP, Yu X, Hermanson S, Gowda K, Brumm P, Mead D, Balan V, Dale BE. Hemicellulases and auxiliary enzymes for improved conversion of lignocellulosic biomass to monosaccharides. Biotechnol Biofuels. 2011;4:5. doi:10.1186/1754-6834-4-5.

Giacobbe S, Balan V, Montella S, Fagnano M, Mori M, Faraco V. Assessment of bacterial and fungal (hemi)cellulose-degrading enzymes in saccharification of ammonia fibre expansion-pretreated Arundo donax. Appl Microbiol Biotechnol. 2015. doi 10.1007/s00253-015-7066-3.

Giacobbe S, Vincent F, Faraco V. Development of an improved variant of GH51 α-L-arabinofuranosidase from *Pleurotus ostreatus* by directed evolution. New Biotechnol. 2014;31(3):230–6.

Guerfali M, Saidi A, Gargouri A, Belghith H. Enhanced enzymatic hydrolysis of waste paper for ethanol production using separate saccharification and fermentation. Appl Biochem Biotechnol. 2015;175:25–42.

Hames B, Ruiz R, Scarlata C, Sluiter A, Sluiter J, Templeton D. Preparation of samples for compositional analysis laboratory analytical procedure (LAP). 2008.

Holtzapple MT, Jun J-H, Ashok G, Patibandla SL, Dale BE. The ammonia freeze explosion (AFEX) process. Appl Biochem Biotechnol. 1991;28–29(1):59–74.

Holtzapple MT, Lundeen JE, Sturgis R, Lewis HE, Dale BE. Pretreatment of lignoceilulosic municipal solid waste by ammonia fiber explosion (AFEX). Appl Biochem Botechnol. 1992;34–35(1):5–21.

Jang YS, Kim B, Shin JH, Choi YJ, Choi S, Song CW, Lee J, Park HG, Lee SY. Bio-based production of C2–C6 platform chemicals. Biotechnol Bioeng. 2012;109(10):2437–59.

Jørgensen H, Kristensen JB, Felby C. Enzymatic conversion of lignocellulose into fermentable sugars: challenges and opportunities. Biofuels Bioprod Biorefin. 2007;1:119–34.

Kim SB, Kim HJ, Kim CJ. Enhancement of the enzymatic digestibility of waste newspaper using Tween. Appl Biochem Biotechnol. 2006;129–132:486–95.

Kuhad RC, Mehta G, Gupta R, Sharma KK. Fed batch enzymatic saccharification of newspaper cellulosics improves the sugar content in the hydrolysates and eventually the ethanol fermentation by *Saccharomyces cerevisiae*. Biomass Bioenergy. 2010;34:1189–94.

Kumar R, Mago G, Balan V, Wyman CE. Physical and chemical characterizations of corn stover and poplar solids resulting from leading pretreatment technologies. Bioresour Technol. 2009;100:3948–62.

Liguori R, Amore A, Faraco V. Waste valorization by biotechnological conversion into added value products. Appl Microbiol Biotechnol. 2013;97(14):6129–47.

Menon V, Rao M. Trends in bioconversion of lignocellulose: biofuels platform chemicals and biorefinery concept. Prog Energy Combust Sci. 2012;38(4):522–50.

Orozco AM, Al-Muhtaseb AH, Rooney D, Walker GM, Aiouache F, Ahmad M. Fermentable sugars recovery from lignocellulosic waste-newspaper by catalytic hydrolysis. Environ Technol. 2013;34:3005–16.

Sangkharak K. Optimization of enzymatic hydrolysis for ethanol production by simultaneous saccharification and fermentation of wastepaper. Waste Manage Res. 2011;29:1134–44.
Sluiter A, Hames B, Hyman D, Payne C, Ruiz R, Scarlata C, Sluiter J, Templeton D, Wolfe J. Determination of total solids in biomass and total dissolved solids in liquid process samples. Laboratory analytical procedure (LAP). 2008a.
Sluiter A, Hames B, Ruiz R, Scarlata C, Sluiter J, and Templeton D. Determination of ash in biomass. Laboratory analytical procedure (LAP). 2008b.
Sluiter A, Hames B, Ruiz R, Scarlata C, Sluiter J, Templeton D, Crocker D. Determination of structural crbohydrates and lignin in biomass. Laboratory analytical procedure (LAPc). 2008c.
Sluiter JB, Ruiz RO, Scarlat CJ, Sluiter AD, Templeton DW. Compositional analysis of lignocellulosic feedstocks. 1. Review and description of methods. J Agric Food Chem. 2010;58(16):9043–53.
Subhedar PB, Babu NR, Gogate PR. Intensification of enzymatic hydrolysis of waste newspaper using ultrasound for fermentable sugar production. Ultrason Sonochem. 2015;22(3):26–32.
Subhedar PB, Gogate PR. Alkaline and ultrasound assisted alkaline pretreatment for intensification of delignification process from sustainable raw-material. Ultrason Sonochem. 2014;21(1):216–25.
Templeton DW, Scarlata CJ, Sluiter JB, Wolfrum EJ. Compositional analysis of lignocellulosic feedstocks. 2. Method uncertainties. J Agric Food Chem. 2010;58(16):9054–62.
Wang L, Sharifzadeh M, Templerb R, Murphy RJ. Technology performance and economic feasibility of bioethanol production from various waste papers. Energy Environ Sci. 2012a;2:5717–30.
Wang L, Sharifzadeh M, Templer R, Murphy RJ. Bioethanol production from various waste papers: economic feasibility and sensitivity analysis. Appl Energy. 2013;111:1172–82.
Wang L, Templer R, Murphy RJ. High-solids loading enzymatic hydrolysis of waste papers for biofuel production. Appl Energy. 2012b;99:23–31.
Wua F-C, Huangb S-S, Shih I-L. Sequential hydrolysis of waste newspaper and bioethanol production from the hydrolysate. Bioresour Technol. 2014;167:159–68.
Wu F-C, Huang S-S, Shih I-L. Sequential hydrolysis of waste newspaper and bioethanol production from the hydrolysate. Bioresour Technol. 2014;167:159–68.
Xin F, Geng A, Chen ML, Gum MJM. Enzymatic hydrolysis of sodium dodecyl sulphate (SDS)-pretreated newspaper for cellulosic ethanol production by *Saccharomyces cerevisiae* and *Pichia stipitis*. Appl Biochem Biotechnol. 2010;162:1052–64.

Chapter 3

Lignocellulosic biomasses as source of biocatalysts

An enormous range of different habitats such as soil, compost piles, decaying plant materials, rumens, sewage sludge, termites gut, forest waste piles, wood processing plants, animal faces, paper mills and hot springs have been so far investigated as sources of both aerobic and anaerobic bacteria producing several different enzymes of industrial interest. In particular, lignocellulosic biomasses represent complex ecosystems in which environmental conditions influence living organisms. As a consequence, autochthonous microbial communities may prevail over other microorganisms due to their ability to produce enzymes involved in the degradation of (hemi)celluloses. Thus, in collaboration with the team of Professor Olimpia Pepe of Department of Agriculture of University of Naples "Federico II", the microbial diversity of natural ecosystems, represented by lignocellulosic biomasses of *Arundo donax, Eucalyptus camaldulensis* and *Populus nigra*, was evaluated by culture-dependent and culture-independent approaches in order to identify novel biocatalysts to overcome costly hurdles due to the intrinsic expense of hydrolytic enzyme production. Chipped wood from *A. donax, E. camaldulensis* and *P. nigra* was processed for biodegradation up to 180 days under two different environmental conditions in order to increase microbial biodiversity: open field and underwood conditions. Samples were collected immediately after preparation (T0) and at 45, 90, 135 and 180 days (T1, T2, T3 and T4, respectively) of biodegradation.

3.1 - Identification of novel (hemi)cellulases by microorganisms cultivation based approach

A culture-dependent approach was carried out in order to isolate putative cellulolytic strains. From samples of the *A. donax, E. camaldulensis,* and *P. nigra* biomasses, the team of Professor Olimpia Pepe isolated and purified 1157 colonies (33% eukaryotes and 67% procariotes) that were assayed all isolates for different enzymatic activities on solid differential media containing specific substrates (such as carboxymethylcellulose - CMC -, Xylan, Pectin and AVICEL) in order to establish the number of putative multifunctional-degrading mcroorganisms.

3.1.1 – Xylanase(s) from ET2C-75A *Pediococcus Acidilactici*

Results and discussion

16S rRNA gene sequencing of the forty bacteria with the highest endo- or exo-cellulase activity on solid was carried out and the identified strains were grouped by phylogenetic analysis into five different clusters, generating a *consensus* tree, as described in Ventorino et al., 2015. Thus, sixteen bacterial strains from previous chosen fourty colonies were selected based on at least one of these characteristics: higher halo dimension on CMC agar, multienzymatic activities, high pectinase activity or high peroxidase activity, and no pathogenicity for humans; these strains were submitted to a quantitative CMCase assay. A further screening of these microorganisms in liquid CMC medium was carried out in order to test the endo-1,4-ß-glucanase activity and the best 6 strains were chosen: PT0-81C - *Xanthomonas Campestris*; ET2C-75A - *Pediococcus acidilactici*; AT1SB-76A - *Pantoea ananatis*; AT1C-83C - *Curtobacterium flaccumfaciens*; AT0-84C - *Curtobacterium sp. citreum*; ET1C-86A - *Promicromonospora Sukumoe.* Based on the literature data reported on these strains (Lednicka et al., 2000; Gagne-Bourgue et

al., 2012; Contreras Sánchez-Matamoros et al., 2013; Hofvendahl and Hahn, 2000), the PT0-81C was excluded from subsequent analyses due to several in-depth studies related to the (hemi)cellulases produced by this strain. The other 5 strains were tested on liquid Avicel and xylan media in order to detect the exo-glucanase and xylanase activity respectively (Table 2).

Table 2. Maximum endo-cellulase, exo-cellulase and xylanase activity of the 5 strains tested on liquid media

	Identification 16S (%)	Maximum activity towards CMC		Maximum activity towards Avicel		Maximum activity towards xylan	
		Value (U/mL)	*Time (h)*	*Value (U/mL)*	*Time (h)*	*Value (U/mL)*	*Time (h)*
ET2C-75A	*Pediococcus acidilactici* (99%)	0.32	12	0.10	18	1.00	72
AT1SB-76A	*Pantoea Ananatis* (99%)	0.18	8	0.10	14	0.10	22
AT1C-83C	*Curctobacterium flaccumfaciens* (99%)	0.20	15	0.20	14	0.28	14
AT0-84C	*Curctobacterium sp. Citreum* (99%)	0.28	17	0.11	10	0.18	24
ET1C-86A	*Promicomonosp. sukumoe* (99%)	0.20	16	0.16	14	0.25	22

All the strains showed an endocellulase activity not exceeding 0.32 U/mL towards CMC and an exo-cellulase activity lower than 0.2 U/mL towards Avicel. However, the ET2C-75A - *Pediococcus acidilactici* strain revealed to be noteworthy because of the highest endo-1,4-ß-glucanase activity (0.32 U/mL after 12 h of incubation) and a remarkable and costant xylanase activity (with a value of 1 U/mL from the 3th to the 9 th day of growth in liquid medium with xylan as unique carbon source). The identification of the protein(s) involved in the hydrolysis of xylan produced by this strain was carried out due to the fact that no xylanases from *Pediococcus acidilactici* was already annotated in databases. In collaboration with the group of Prof. Birolo, Department of Chemical Science, University of Naples "Federico II", a zymographic approach combined with proteomic analysis was performed. The figure 4 shows the semi-denaturing Sodium Dodecyl Sulphate - PolyAcrylamide Gel Electrophoresis (SDS-PAGE) of the culture supernatant from the sample ET2C-75A after 3 days of growth. After protein preparation, the lane 1 was incubated on agar medium containing xylan as carbon source and tested for xylanase activity by Congo Red assay. Two halos of activity were detected. The lanes 2 (protein marker) and 3 (sample) were stained by Coomassie Brilliant Blue: the highlighted areas show the protein bands matching the halos of xylanase activity in lane 1.

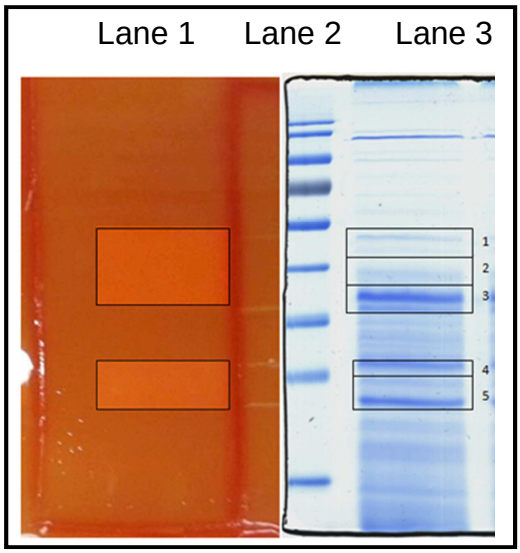


Figure 4. Zymography and semi-denaturing SDS-PAGE of proteins in supernatant from the sample ET2C-75A (3th growth day): lane 1 = Congo Red activity assay - the highlighted areas show the halos of xylanase activity; lane 2= protein marker; lane 3 = Coomassie Brilliant Blue staining - the highlighted areas show the protein bands matching the halo of xylanase activity in lane 1.

The five different commassie stained bands of proteins matching the halo of xylanase activity by Congo Red were excised from the gel and an in-situ-hydrolysis was performed using trypsin as protease. The peptide mixture was filtered and subjected to desalting by chromatography with C18 resins in order to perform analysis by tandem mass spectrometry (LC-MS / MS). The analysis data were matched with the NCBInr database limited to the taxonomic kingdom bacteria by using Mascot software. In particular, the proteins excised from the bands 3 and 4 (Table 3), corresponding to the first and the second activity positive halo, matched three endo-1,4-beta-xylanases (from *Cellulomonas flavigena* DSM 20109, *Cellulomonas flavigena* DSM 20109* and *Actinosynnema mirum* DSM 43827) and one β-1,4-xylanase (from *Actinoplanes sp.* SE50/110), involved in the hydrolysis of beta-1,4-xylan into xylose.

Table 3. Data by analysis LC – MS/MS of selected 3 and 4 bands. For each band is shown: the identification code of the protein(s) matched in NCBInr database; the name of the identified protein(s) and the species of origin; the nominal mass of the protein(s); the score assigned by Mascot software; the number of peptides identified; the coverage of the allignment.

Band	NCBInr ID	Protein	nominal mass (Mr)	Score	n° peptides	Sequence coverage %
3	gi|334338003	xyloseisomerase [Isoptericolavariabilis 225]	42736	74	3	11
	gi|296131352	endo-1,4-beta-xylanase [Cellulomonasflavigena DSM 20109]	50621	64	3	6
	gi|296130324	endo-1,4-beta-xylanase [Cellulomonasflavigena DSM 20109]	87951	55	3	4
	gi|386847953	Beta-1,4-xylanase [Actinoplanes sp. SE50/110]	85638	49	2	3
4	gi|296131352	endo-1,4-beta-xylanase [Cellulomonasflavigena DSM 20109]	50621	208	4	10
	gi|296130324	endo-1,4-beta-xylanase [Cellulomonasflavigena DSM 20109]	87951	177	3	5
	gi|334338003	Xyloseisomerase [Isoptericolavariabilis 225]	42736	131	6	15
	gi|256376709	endo-1,4-beta-xylanase [Actinosynnemamirum DSM 43827]	47677	130	2	7

	gi|254785068	kinase, pfkB family [Teredinibacterturnerae T7901]	39870	72	2	4
	gi|152967235	glutamine synthetase, type I [Kineococcusradiotolerans SRS30216]	53635	49	2	4

The matching of the excised peptides with sequences of the four above mentioned enzymes (figure 5) suggests the analogy between these xylanases and the xylanase(s) from *Pediococcus acidilactici*.

Figure 5. Alignment of the sequences of endo-1,4-βxylanases by using ClustalW. Each sequence is identified by UniProt code (D5UDE7: protein from *Cellulomonas flavigena* DSM 20109; D5UI30: protein from *Cellulomonas flavigena* DSM 20109*; C6WMJ4: protein from *Actinosynnema mirum* DSM 43827; G8S9Z5: protein from *Actinoplanes sp.* SE50/110). The presence of the same amino acid in all the sequences are shown by *; the symbols : and . idicate greater or lesser difference between the amino acids in the same position in the sequences. The highlighted peptides match with peptides from the excised commassie stained bands.

The fact that none of the identified proteins is from *Pediococcus acidilactici* can be explained with the partial annotation of the so far available sequenced genomes of different strains of *P. acidilactici*. These data encourage further researches on this bacterial strain with the aim of the identification and characterization of novel biocatalysts involved in the degradation of (hemi)cellulose. Although fungi play a key role in the degradation of plant materials, nowadays, the isolation of bacteria as a source of novel enzymes for cellulose hydrolysis is strongly investigated. The several advantages exhibited by bacteria strains in comparison to fungi are related to the high growth rates, the possibility to easily engineer them, and, last but not least, the higher probability to isolate them from environmental niches that naturally drive them to the production of enzymes resistant to environmental stresses.

Materials and methods

- The lignocellulosic biomasses and sampling, the screening for functional activities on solid media and the identification and phylogenetic analysis of cellulolytic strains were desrcribed in detail in Ventorino et al., 2015

- Screening of endo-1,4-ß-glucanase, exo-glucanase and endo-xylanase activity in liquid media

The bacterial strains were pre-inoculated by dissolving a single colony in 3 mL of liquid medium containing, 0.7% yeast extract, 4 g L^{-1} KH_2PO_4, 4 g L^{-1} Na_2HPO_4, 0.2 g L^{-1} $MgSO_4 \cdot 7H_2O$, 0.001 g L^{-1} $CaCl_2 \cdot 2H_2O$ and 0.004 g L^{-1} $FeSO_4 \cdot 7H_2O$. 0.5% CMC, Avicel or xylan were added to liquid medium for endo-1,4-ß-glucanase, exo-glucanase and endo-xylanase screening respectively. After overnight incubation at 37°C, a volume of the broth culture corresponding to 0.1 O.D. was used to inoculate 100-mL plugged Erlenmeyer flasks, each containing 20 mL of the same medium. During incubation at 37°C on a rotary shaker at 225 rpm, samples of the liquid culture were withdrawn and used to measure the optical density ($O.D._{600nm}$). The extracellular *endo*-1,4-ß-glucanase activity was assayed by AZO-CMCase assay (Megazyme, Ireland), following the supplier's instructions. The extracellular exo-glucanase activity was assayed by the 3,5-dinitrosalicylic acid (DNS) method, according to Wood et al., 1988. The extracellular endo-xylanase activity was assayed by the DNS method, according to Bailey et al., 1992. The analytical determinations correspond to the mean value of three replicates.

- Determination of protein concentration

Protein concentration of crude enzyme preparation was determined by Bradford method using Biorad reactive (München, Germany) following the procedure suggested by the supplier. Bovin serum albumin (BSA) was used to set up the standard curve.

- Zymogram analyses

Semi-denaturing gel electrophoresis was carried out loading non-denatured and not-reduced samples on a SDS polyacrylamide gel, performed as described by Laemmli, 1970. Proteins showing xylanase activity were visualized following a modified version of the assay reported by Béguin, 1983. After electrophoresis, the gel was soaked in the same buffer used for dissolving proteins and gently shaken to remove SDS and renature the proteins in the gel. The gel was then laid on the top of a thin sheet of 1.5% agar containing 1% xylan. After 2 h incubation at 50°C, halo of xylan hydrolysis were revealed by staining the agar replica with 0.1% of Congo red (30 minutes of Congo red incubation at room temperature followed by washing with 5 M NaCl).

- Protein identification by mass spectrometry

Slices of interest from the semi-denaturing PAGE were cut and *in situ* digested after extensive destaining with 0.1 M NH_4HCO_3 pH 7.5 and acetonitrile, reduction of disulphide bonds for 45 minutes in 100 µl of 10 mM dithiothreitol, 0.1 M NH_4HCO_3, pH 7.5 and carboxyamidomethylation of thiols for 30 minutes in the dark by addition of 100 µl of 55 mM iodoacetamide dissolved in the same buffer. Enzymatic digestion was performed by adding to each slice 100 ng of proteomic-grade trypsin in 10 µl of

10 mM NH_4HCO_3 pH 7.5 for 2 hours at 4°C. The buffer solution was then removed and 50 µl of 10 mM NH_4HCO_3 pH 7.5 were added and incubated for 18 hours at 37°C. Peptides were extracted with 20 µl of 10 mM NH_4HCO_3, 1% formic acid, 50% acetonitrile at room temperature. Peptide mixtures were filtered on 0.22 µm PVDF membrane (Millipore) and analysed by LC-MS/MS on a 6520 Accurate-Mass Q-TOF LC/MS System (Agilent Technologies, Palo Alto, CA) equipped with a 1200 HPLC system and a chip cube (Agilent Technologies). After loading, the peptide mixture is concentrated and washed in 40 nL enrichment column (Agilent Technologies chip), with 0.1% formic acid in 2% acetonitrile as the eluent. The sample is then fractionated on a C18 reverse-phase capillary column (Agilent Technologies chip) at a flow rate of 400 nl/min, with a linear gradient of eluent B (0.1% formic acid in 95% acetonitrile) in A (0.1% formic acid in 2% acetonitrile) from 7 to 80% in 50 min. Peptide analysis is performed using data-dependent acquisition of one MS scan (mass range from 300 to 1800 m/z) followed by MS/MS scans of the five most abundant ions in each MS scan. MS/MS spectra were measured automatically when the MS signal is over the threshold of 50000 counts. Double and triple charged ions were preferably isolated and fragmented over single charged ions. Raw data from nanoLC–MS/MS analyses transformed in *mz.data* format and used to query non-redundant protein databases with a licensed version of MASCOT 2.1 (Matrix Science, Boston, USA). Additional search parameters were a peptide mass tolerance set at 10 ppm and a fragment mass tolerance of 0.6 Da, up to 3 allowed missed cleavages, carbamidomethylation of cysteines as fixed modification, oxidation of methionine, and cyclization of N-term Q to pyro-Glu as variable modifications. Only doubly and triply charge ions were considered. Ions score is -10 log(P), where P is the probability that the observed match is a random event. The threshold above which the individual ions score indicates identity or extensive homology ($p < 0.05$) can vary from search to search. In our searches, on average, individual ion scores >25 indicated identity or extensive homology ($p < 0.05$). Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits (http://www.matrixscience.com/help/interpretation_help.html). Trypsin, dithiothreitol, iodoacetamide and NH_4HCO_3 were purchased from Sigma. Trifluoroacetic acid (TFA)-HPLC grade was from Carlo Erba (Milan, Italy). All other reagents and solvents were of the highest purity available from Baker.

3.1.2 – Endo-cellulase(s) from AE-T0-58P *Streptomyces sp*.

Results and discussion

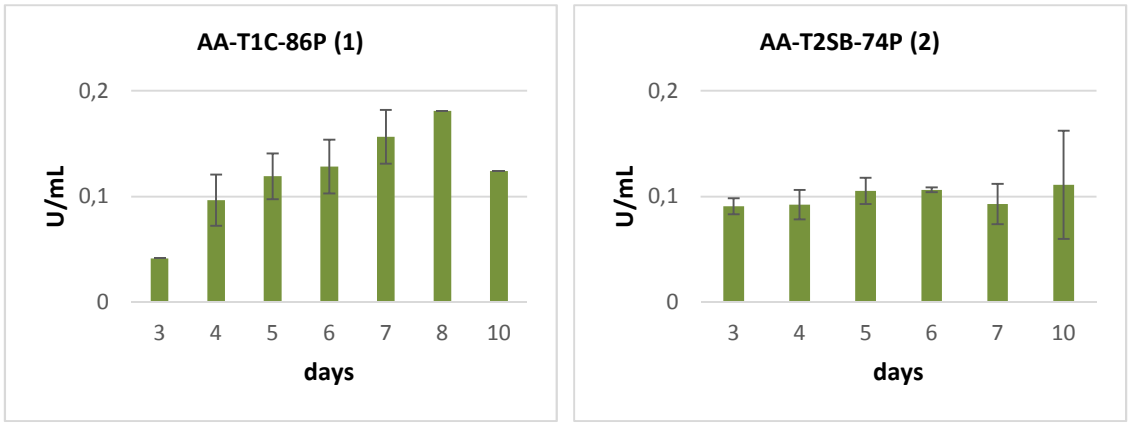
A further screening of 1157 colonies isolated and purified by the team of Professor Olimpia Pepe from biodeteriorated *A. donax, E. camaldulensis,* and *P. nigra* biomasses was performed in order to detect Actinomycetes strains with lignocellulose-degradation potentiality.

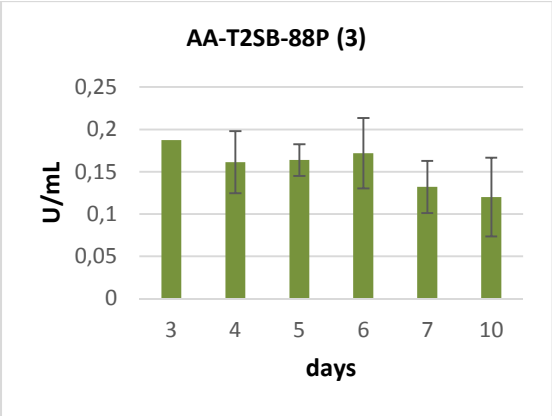
Twenty-four strains were selected and phenotypically characterized by analysis of colony and cell morphology, and Gram reaction. The identification by 16S rRNA gene sequencing showed that all the strains belonged to the *Streptomyces* genus. Moreover, an activity halo on CMC solid medium with a diameter ≥ 20 mm was produced by all 24 selected microorganisms (Table 4).

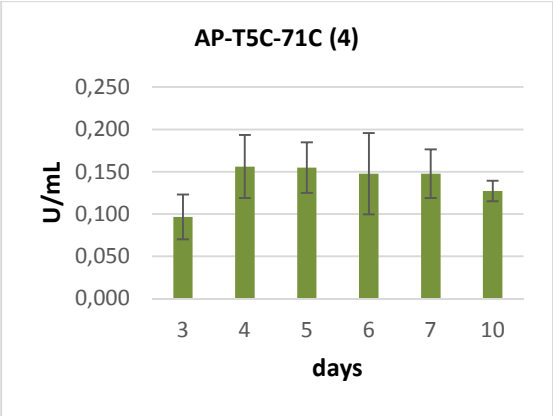
Table 4 - CMC activity halo (mm) and 16S identification of the 24 best endo-cellulase(s) producers belonged to *Streptomyces* genus.

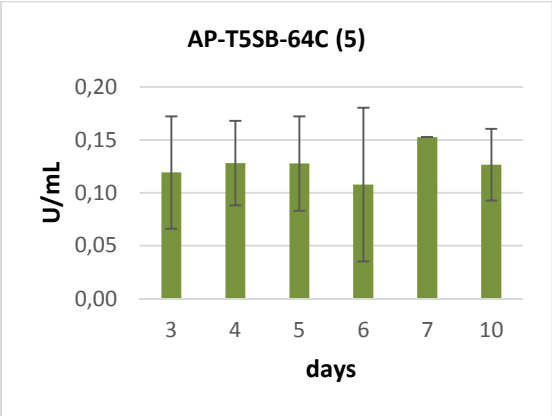
Strain	CMC activity halo	Identification 16 S (% Identity)
AA-T1C-86P (1)	28 mm	*Streptomyces drozdowiczii (99%)*
AA-T2SB-74P (2)	30 mm	*Streptomyces lienomycini (99%)*
AA-T2SB-88P (3)	30 mm	*Streptomyces flavogriseus/ Streptomyces flavovirens (99%)*
AA-T2SB-74P (4)	26 mm	*Streptomyces flavogriseus/ Streptomyces flavovirens (99%)*
AP-T5SB-64C (5)	20 mm	*Streptomyces argenteolus (99%)*
AA-T5SB-62C (6)	28 mm	*Streptomyces flavogriseus/ Streptomyces flavovirens (99%)*
AE-T0-53P (7)	30 mm	*Streptomyces flavogriseus/ Streptomyces flavovirens (99%)*
AE-T1SB-73P (8)	30 mm	*Streptomyces fimicarius (99%)*
AE-T2C-612P (9)	30 mm	*Streptomyces flavogriseus/ Streptomyces flavovirens (99%)*
AE-T0-58P (10)	40 mm	*Streptomyces flavogriseus/ Streptomyces flavovirens (99%)*
AE-T2SB-82P (11)	32 mm	*Streptomyces fimicarius (99%)*
AE-T2SB-78P (12)	28 mm	*Streptomyces flavogriseus/ Streptomyces flavovirens (99%)*
AA-T55B-63A (13)	30 mm	*Streptomyces flavogriseus/ Streptomyces flavovirens (99%)*
AE-TO-61P (14)	30 mm	*Streptomyces flavogriseus/ Streptomyces flavovirens (99%)*
AA-T5SB-64C (15)	30 mm	*Streptomyces flavogriseus/ Streptomyces flavovirens (99%)*
AA-T2C-85A (16)	32 mm	*Streptomyces flavogriseus/ Streptomyces flavovirens (99%)*
AP-T1C-71X (17)	36 mm	*Streptomyces albolongus (99%)*
AP-755B-51A (18)	32 mm	*Streptomyces flavogriseus/ Streptomyces flavovirens (99%)*
AE-T1C-618X (19)	30 mm	*Streptomyces fimicarius (99%)*
AP-T1C-73A (20)	24 mm	*Streptomyces flavogriseus/ Streptomyces flavovirens (99%)*
AE-T1SB-710A (21)	36 mm	*Streptomyces flavogriseus/ Streptomyces flavovirens (99%)*
AE-T2C-66P (22)	30 mm	*Streptomyces ambofaciens (98%)*
AA-T1SB-72A (23)	24 mm	*Streptomyces flavogriseus/ Streptomyces flavovirens (99%)*
AA-T25B-714A (24)	26 mm	*Streptomyces flavogriseus/ Streptomyces flavovirens (99%)*

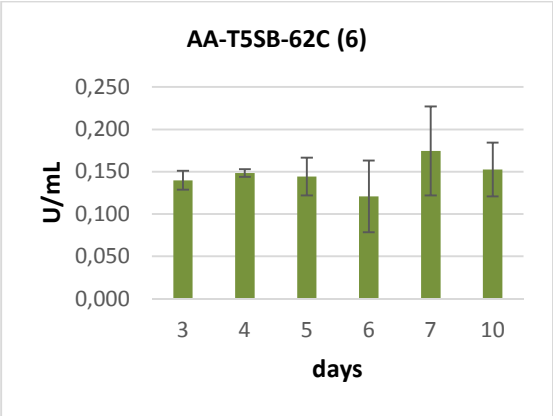
A further screening of these chosen strains in liquid CMC medium was carried out in order to test the endo-1,4-ß-glucanase production and the results are showed in figure 6.

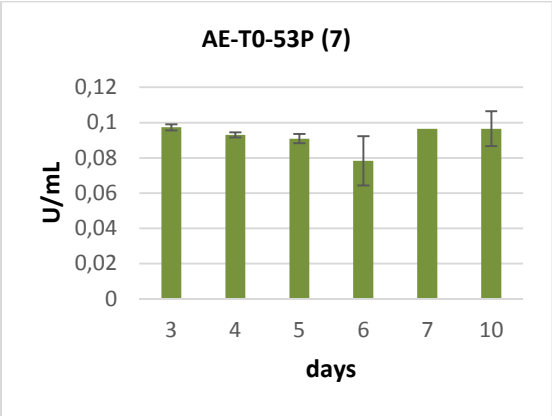


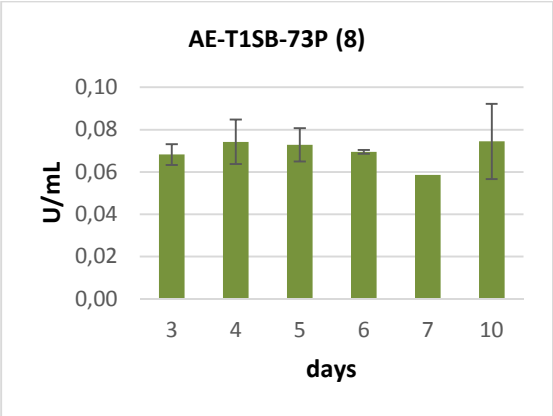

AA-T2SB-88P (3)
U/mL
0,25
0,2
0,15
0,1
0,05
0
3 4 5 6 7 10
days

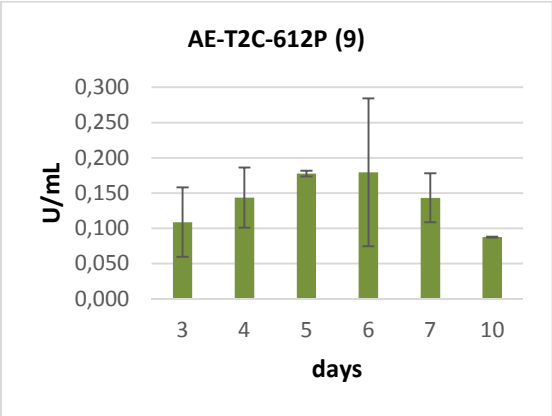

AP-T5C-71C (4)
U/mL
0,250
0,200
0,150
0,100
0,050
0,000
3 4 5 6 7 10
days

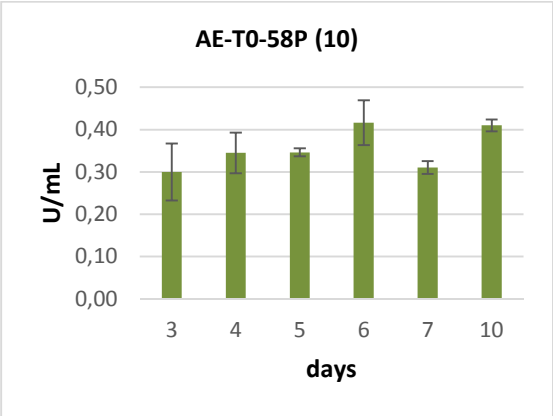

AP-T5SB-64C (5)
U/mL
0,20
0,15
0,10
0,05
0,00
3 4 5 6 7 10
days

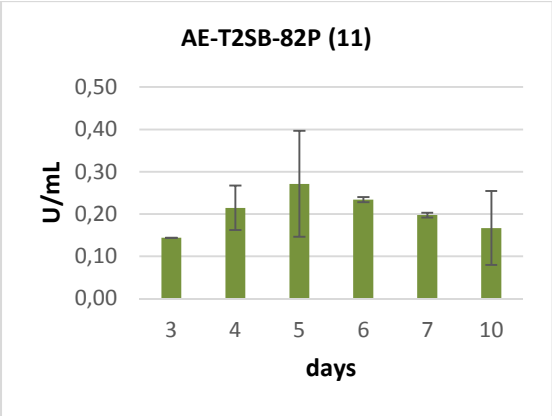

AA-T5SB-62C (6)
U/mL
0,250
0,200
0,150
0,100
0,050
0,000
3 4 5 6 7 10
days

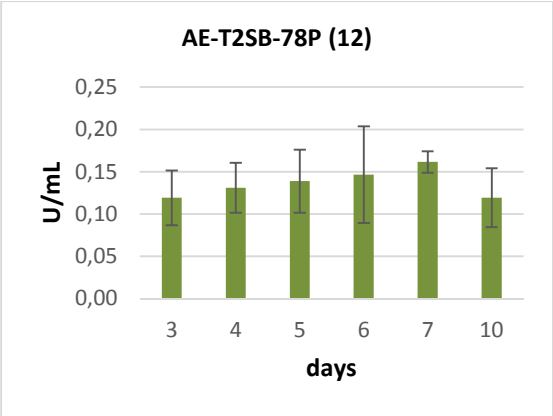

AE-T0-53P (7)
U/mL
0,12
0,1
0,08
0,06
0,04
0,02
0
3 4 5 6 7 10
days

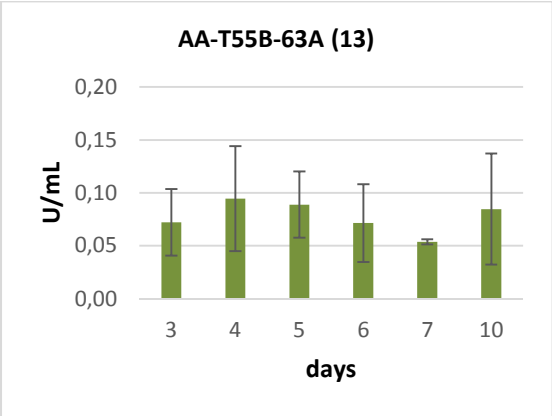

AE-T1SB-73P (8)
U/mL
0,10
0,08
0,06
0,04
0,02
0,00
3 4 5 6 7 10
days

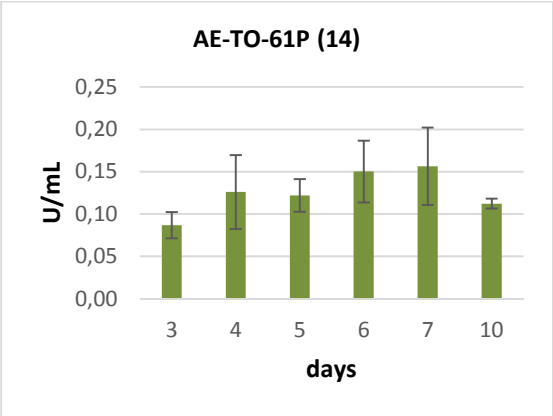

AE-T2C-612P (9)
U/mL
0,300
0,250
0,200
0,150
0,100
0,050
0,000
3 4 5 6 7 10
days

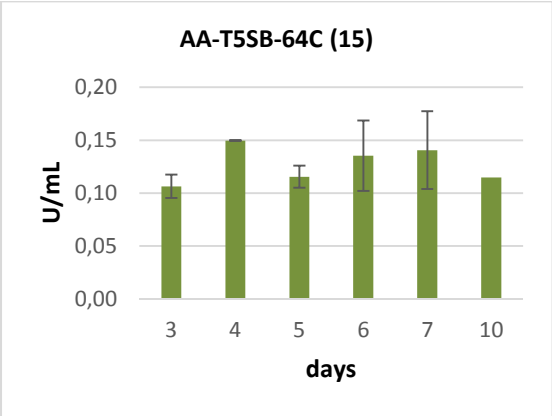

AE-T0-58P (10)
U/mL
0,50
0,40
0,30
0,20
0,10
0,00
3 4 5 6 7 10
days

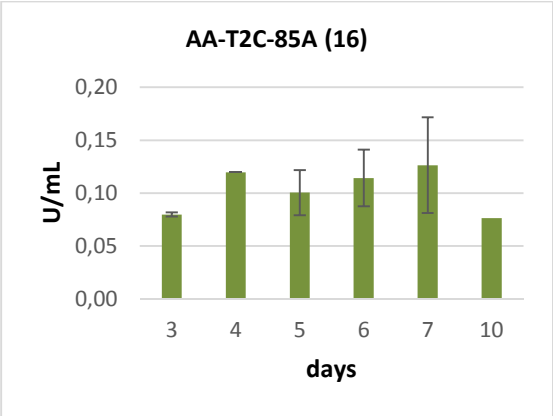

AE-T2SB-82P (11)
0,50
0,40
0,30
0,20
0,10
0,00
U/mL
3
4
5
6
7
10
days

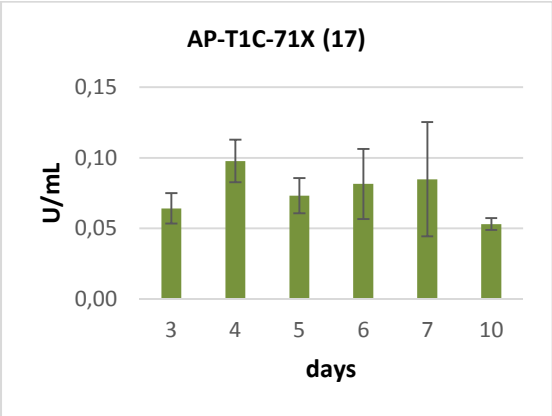

AE-T2SB-78P (12)
0,25
0,20
0,15
0,10
0,05
0,00
U/mL
3
4
5
6
7
10
days

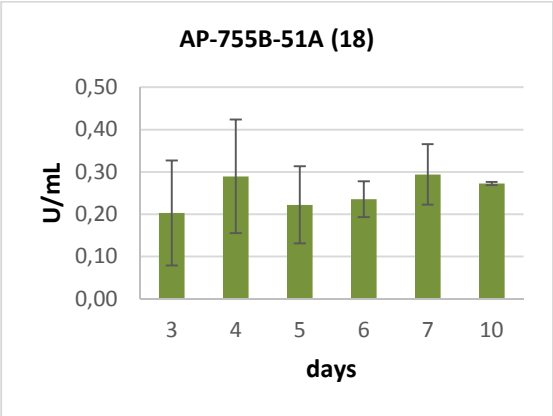

AA-T55B-63A (13)
0,20
0,15
0,10
0,05
0,00
U/mL
3
4
5
6
7
10
days


AE-TO-61P (14)
0,25
0,20
0,15
0,10
0,05
0,00
U/mL
3
4
5
6
7
10
days


AA-T5SB-64C (15)
0,20
0,15
0,10
0,05
0,00
U/mL
3
4
5
6
7
10
days


AA-T2C-85A (16)
0,20
0,15
0,10
0,05
0,00
U/mL
3
4
5
6
7
10
days


AP-T1C-71X (17)
0,15
0,10
0,05
0,00
U/mL
3
4
5
6
7
10
days


AP-755B-51A (18)
0,50
0,40
0,30
0,20
0,10
0,00
U/mL
3
4
5
6
7
10
days

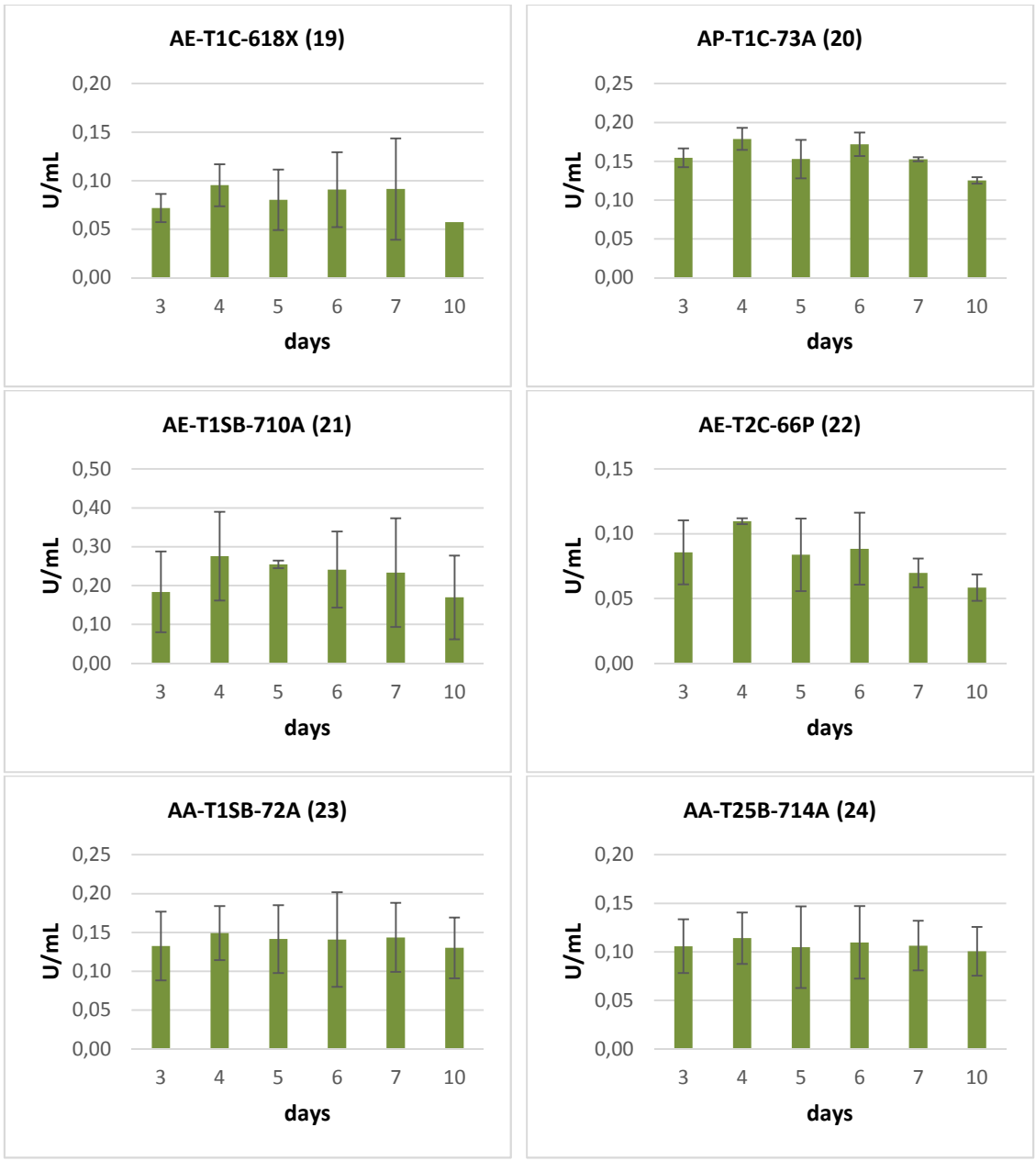


Figure 6 – Time course of endo-cellulase activity production by the 24 selected *Actinomycetes.*

The AE-T0-58P (10) strain produced the highest endo-1,4-ß-glucanase activity level (0.41±0.05 U/mL) and was chosen as reference. The maximum values of endo-cellulase production for AP-755B-51A (18), AE-T1SB-710A (21), AE-T2SB-829 (11) and AE-T2C-612P (9) were at least the 50% in comparison to the maximum value showed by AE-T0-58P (10). These strains were classified as high producers. The strains AP-T1C-73A (20), AA-T2SB-889 (3), AA-T5SB-62C (6), AA-T1C-869 (1), AP-T5C-71C (4), AE-TO-61P (14), AA-T1SB-72A (23), AE-T2SB-78P (12) showed a maximum value of activity between 34 and 50 % in comparison to the reference value. The other strains produced an endo-cellulase activity less than the 34% of the reference.

In order to evaluate the best conditions for a long-term storage, the residual enzymatic activity of the best five producers were assayed at 4°C, -20°C and -80°C up to 30 days of storage. The AE-T1SB-710A (21) strain retained most of the endo-1,4-ß-glucanase activity after 30 days (60% at 4°C and more than 80% at -20°C

and -80°C); the AE-T0-58P (10) strain preserved 50-60% enzymatic activity after 30 days at all three temperatures; the AE-T2SB-82P (11) and AP-755B-51A (18) strains retained 80-100% of activity at -20° and -80°C up to 30 days and lost 50% activity after 6 days at 4°C; contrariwise no selected storage conditions were suitable for the endo-cellulase(s) produced by the AE-T2C-612P (9) strain (data not shown). Based on these results, the AE-T2C-612P (9) strain was not subjected to further experiments. The optimal temperature and pH and thermoresistance of endo-cellulase activities were also evaluated.

In the Na-citrate buffer, the cellulase activity produced by the analyzed *Streptomyces* strains showed an optimum at same pH value of 5.0 (figure 7). Further the optimum of temperature for all the strains was at 50°C (figure 8), which represents the condition mostly used for the enzymatic hydrolysis of lignocellulosic biomasses.

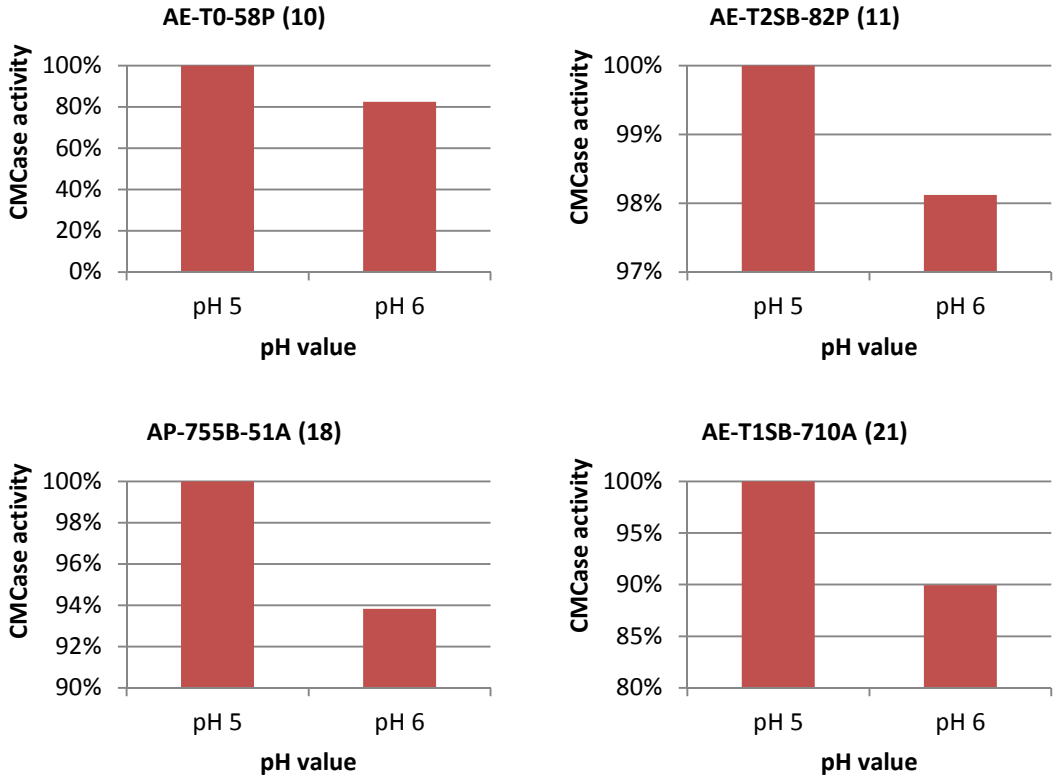


Figure 7 – Effect of pH in 50 mM Na-citrate buffer on endo-cellulase activity produced by the 4 selected strains.

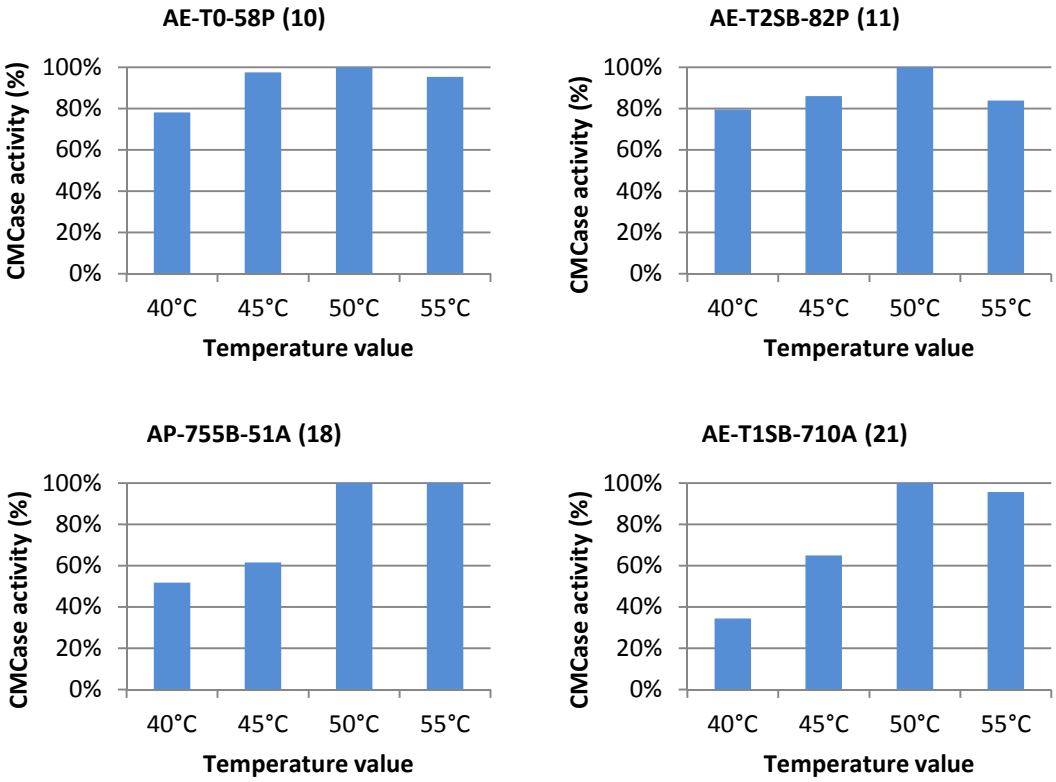
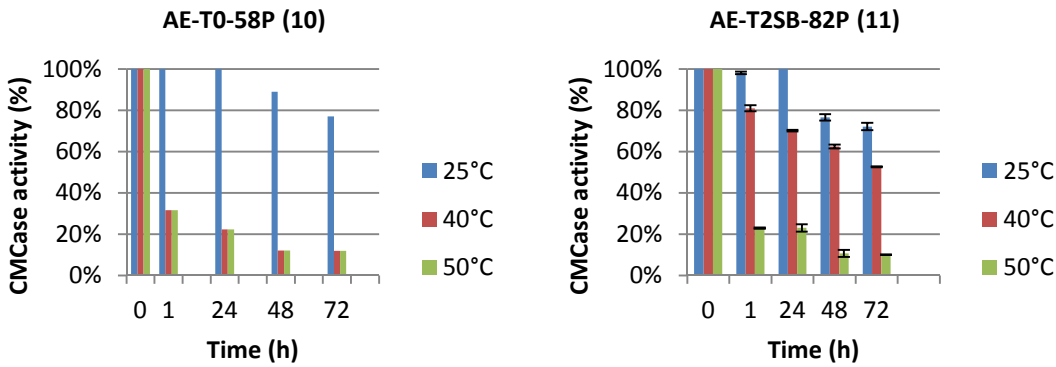


Figure 8 – Effect of temperature on endo-cellulase activity produced by the 4 selected strains.

As shown in figure 9, all the tested cellulases showed very low stability at 50°C losing more than 50% of activity after only 1 hour. The enzyme produced by the AE-T0-58P (10) strain was the least stable at 40°C (35% residual activity after 1 hour), whereas the cellulase(s) from other strains retained more than 50% activity after 72 hours. All the tested enzymes were highly stable at 25°C: in particular the AE-T0-58P (10), AE-T2SB-82P (11), AP-755B-51A (18) and AE-T1SB-710A (21) retained respectively 77%, 72%, 82% and 71% of initial activity.



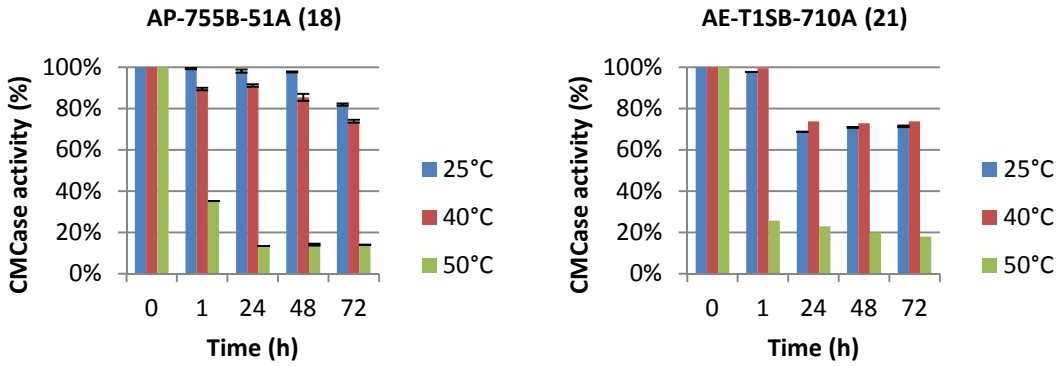


Figure 9 – Thermoresistence of the endo-cellulase activity produced by the 4 selected strains.

The supernatants from the four endo-cellulase(s) producer strains were tested in the hydrolysis of pretreated *A. donax*.

The saccharification experiments were carried out by using a mix of enzymes due to the fact that, as shown in several papers, a mix of (hemi)cellulases is crucial to enhance hydrolysis of pretreated lignocellulosic biomasses. An enzyme cocktails prepared with the commercial products Accellerase 1500, Accellerase XY, and Accellerase BG (mix 1) was chosen as reference mix. In order to evaluate the hydrolysis yield by using the endo-cellulase(s) produced by AE-T0-58P (10), AE-T2SB-82P (11), AP-755B-51A (18) and AE-T1SB-710A (21) strains in substitution of the corresponding commercial enzyme, the supernatants culture were used to replace the Accellerase 1500 in the mix 1, obtaining mix 2, mix 3, mix 4 and mix 5 respectively. The chosen conditions of pH and temperature adopted in saccharification experiments were those at which the tested enzyme(s) retained 80% of the initial activity. Table 5 summarizes the conversion results of pretreated *A. donax* biomass using the different enzymatic mixtures. Although the hydrolysis yields obtained by using mixes 3, 4 and 5 were very low, after 72 h hydrolysis the substitution of Accellerase 1500 with the supernatant containing cellulase(s) from AE-T0-58P (10) strain (mix 2) allowed to obtain a glucose and xylose yield of 4.47±0.5 g/L and 5.87±0.2 g/L respectively. Interestingly, these yields were 82% and 85% respectively of the corresponding values obtained by using the reference commercial mixes (mix 1). This comparison showed that the endo-cellulase activity produced by the AE-T0-58P (10) strain was a good candidate to replace the commercial cellulose mix Accelerase 1500 for the *A. donax* saccharification with satisfactory conversion yields.

Table 5 – Glucose and xylose yields of the saccharification of pretreated *A. donax* by using mix 1 (Accelerase 1500, Accellerase XY and Accellerase BG), mix 2 (supernatant from AE-T0-58P (10), Accellerase XY and Accellerase BG), mix 3 (supernatant from AE-T2SB-829 (11), Accellerase XY and Accellerase BG), mix 4 (supernatant from AP-755B-51A (18), Accellerase XY and Accellerase BG) and mix 5 (supernatant from AE-T1SB-710A (21), Accellerase XY and Accellerase BG). The percentage yields are referred to glucose and xylose yield (100%) by using reference mix 1.

Mix	Glucose Yield (g/L)				Xylose Yield (g/L)			
	48 h		*72 h*		*48 h*		*72 h*	
mix 1	4.54 ± 0.28	100%	5.43 ± 0.15	100%	3.92 ± 0.89	100%	6.89 ± 0.54	100%
mix 2	2.40 ± 0.2	53%	4.47 ± 0.5	82%	3.4 ± 0.2	87%	5.87 ± 0.2	85%
mix 3	1.70 ± 0.3	37%	2.2 ± 0.84	40%	1.8 ± 0.4	46%	2.87 ± 1.25	42%
mix 4	1.71 ± 0.1	37%	2.11 ± 0.07	39%	1.95 ± 0.07	50%	2.5 ± 0.1	36%
mix 5	1.88 ± 0.16	41%	1.98 ± 0.39	36%	2.85 ± 0.4	73%	2.87 ± 0.39	42%

Actinomycetes, usually isolated from soils, represent one of the most diverse groups of filamentous bacteria capable of surviving in a number of ecological niches due to their bioactive potential and have gained special importance mainly as the most potent source of antibiotics (Kandasamy et al., 2012) and other bioactive secondary metabolites (Solecka et al., 2012) useful in several biotechnological and medical areas. On the other hand, only few reports have focused on potential of Actinomycetes as source of enzymes, although they are considered a promising source of a wide range of biocatalysts. The encouraging data obtained in this study confirm the potentiality of many strains of *Streptomyces* genus to produce enzymes useful for the bioconversion of underutilized agricultural and urban wastes into high-value chemical products (Crawford, 1988) due to their ability to synthesize biocatalysts necessary to allow a complete hydrolysis of the recalcitrant components of the lignocellulosic biomass (Amore et al., 2012a).

Materials and methods

- The lignocellulosic biomasses and sampling, the screening for functional activities on solid media and the identification and phylogenetic analysis of cellulolytic strains were described in detail in Ventorino et al., 2015.

- Screening of endo-1,4-ß-glucanase activity in liquid media

The Actinomycete were maintained on the PCA plates (0.5% tryptone, 0.25% yeast extract, 0.1% glucose and 1.5% agar) and incubated at 28°C. The preinoculum was performed adding one agar plug (4 mm) from two weeks-old plate in 10mL of liquid medium containing CMC (10 gL^{-1} CMC, 0.5 gL^{-1} yeast extract, 2.5 gL^{-1} $(NH_4)_2SO_4$, 2.7 gL^{-1} KH_2PO_4, 5.3 gL^{-1} NA_2HPO_4, 0.2 gL^{-1} NaCl, 0.2 gL^{-1} $MgSO_4$ and 0.05 gL^{-1} $CaCl_2$). The preinoculum was incubated for 48 hours on a rotary shaker (250rmp) at 28°C. A quantity of 2 x 10^6 spores mL^{-1} were inoculated in 100 ml plugged Erlenmeyer flasks in duplicate, each containing 20 mL of the same medium used for the preinoculum. The flasks were incubated for ten days at 28°C on a rotary shaker (200rpm). The endo-1,4-ß-glucanase activity was assayed for ten days every 24 hours by using Azo-CMC (Megazyme, Ireland) as substrate, following supplier's instructions and determined by referring to a standard curve.

- Residual endo-1,4-ß-glucanase activity

The residual enzymatic activity was determined on the supernatant of the *Streptomyces* cultures stored at 4°C, -20°C and -80°C. The endo-1,4-ß-glucanase activity was assayed by using Azo-CMC (Megazyme, Ireland).

- Optimum pH

The optimum pH of cellulases was determined on the supernatant of the *Streptomyces* cultures concentrated by ultrafiltration with a 10 kDa polyethersulfone membrane (Millipore Corporation, Bedford, MA, USA). The experiments were performed at 40°C in 50 mM Na-citrate buffer, at pH 5.0 and 6.0, using Azo-CMC (Megazyme, Ireland) as substrate dissolved in the above-mentioned buffer, following supplier's instructions.

The reported results correspond to mean values of the three independent experiments each one performed in three replicates.

- Optimum temperature and thermoresistance

The supernatant of the cultures was concentrated by ultrafiltration with a 10 kDa polyethersulfone membrane (Millipore Corporation, Bedford, MA, USA) and subjected to the determination of optimum temperature and thermoresistance of the cellulases. To assess the optimum temperature, the substrate Azo-CMC (Megazyme, Ireland) was dissolved in 50 mM Na citrate at pH 5.0 and incubated in presence with the enzymatic preparation at 40, 45, 50 and 55°C.

The thermoresistance of the cellulase activity was investigated by incubating the *Streptomycetes* culture supernatant in 50 mM Na citrate pH 5.0, at 25°C, 40°C, and 50°C.

The reported results correspond to mean values of the three independent experiments, each one performed in three replicates.

- Enzymatic Hydrolysis

The saccharification experiments were carried out at 50 °C for 72 h in a total volume of 5 mL (50 mM Na-citrate buffer pH 5.0 plus enzyme cocktail), at a solid loading of 5%. The hydrolysis of pretreated lignocellulosic materials was carried out with enzyme cocktails prepared with the commercial products Accellerase 1500 (cellulase enzyme complex), Accellerase XY (xylanase enzyme complex), and Accellerase BG (β-glucosidase) provided by Danisco US Inc., Genencor Division (Palo Alto, CA, USA) at the indicated amounts expressed as units per grams of pretreated dry biomass: 5.4 U g−1 of Accellerase 1500, 4000 U g−1 of Accellerase XY and 145 U g−1 of Accellerase BG (mix 1). Moreover, the supernatants from AE-T0-58P (10), AE-T2SB-829 (11), AP-755B-51A (18) and AE-T1SB-710A (21) culture growth were used in substitution of Accellerase 1500 at the same amount of units per grams of pretreated dry biomass. The mixes were named mix 2 (supernatant from AE-T0-58P (10), Accellerase XY and Accellerase BG), mix 3 (supernatant from AE-T2SB-829 (11), Accellerase XY and Accellerase BG), mix 4 (supernatant from AP-755B-51A (18), Accellerase XY and Accellerase BG) and mix 5 (supernatant from AE-T1SB-710A (21), Accellerase XY and Accellerase BG).

The saccharification mixtures together with blanks (pretreated lignocellulosic materials without enzyme cocktail) were incubated in a shaking water bath at 50 °C and 650 rpm for 72 h. Samples were withdrawn at different time intervals, chilled on ice and centrifuged at 16,500 × g for 30 min at 4 °C. The supernatants were analyzed to quantify the amount of sugars released as described below. The saccharification yield was expressed as sugar production calculated respect to the amount of sugars present in the dry weight pretreated materials before the hydrolysis processes. Each saccharification experiment was run in triplicate.

- Determination of Sugar Content

For estimation of the sugars released from enzymatic hydrolysates, the cleared supernatants were opportunely diluted, and analyzed by a high-performance liquid chromatographic (HPLC) system (Dionex, Sunnyvale, CA, USA), equipped with an anionic exchange column (Carbopac PA-100) and a pulsed electrochemical detector. Glucose and xylose were separated with 16 mM sodium hydroxide at a flow rate of 0.25 mL/min, and identified by the respective standards. Fucose was used as internal standard.

3.2 - Identification of novel (hemi)cellulases by metagenomic approach

As the genomic era has shown that only few microorganisms can be cultured under standard laboratory conditions (Whitman et al., 1998), the extraction and analysis of genetic material directly from environmental samples (environmental DNA – eDNA), namely metagenomics, is a promising way to overcome this bottleneck (Sharma et al., 2008).

The metagenomic approach is a multi-step process including the preparation/pretreatment of the samples and the nucleic acids extraction. In the pretreatment step, an enrichment of the sample in target microorganisms can be performed to improve the specificity of the DNA samples; non-enrichment of samples can be used to preserve the natural diversity and relative abundance of microbial communities. Two different screening methodologies can be used on metagenomic material: function-driven approach of expression libraries or sequence-driven analysis. The latter one can be achieved by either an homology-based screening of libraries (using engineered degenerate primers from highly conserved target regions in PCR or hybridization) or the direct large-scale sequencing of the extracted eDNA. The large-scale sequencing is carried out by shotgun-sequencing (the extraction of a large amount of eDNA, its random shearing into smaller fragments and direct sequencing) or Amplicon sequencing (the filtering of target gene fragments by amplification of an informative marker by Polymerase Chain Reaction - PCR). In the figure 10 an overview of the possible metagenomic strategies is reported.

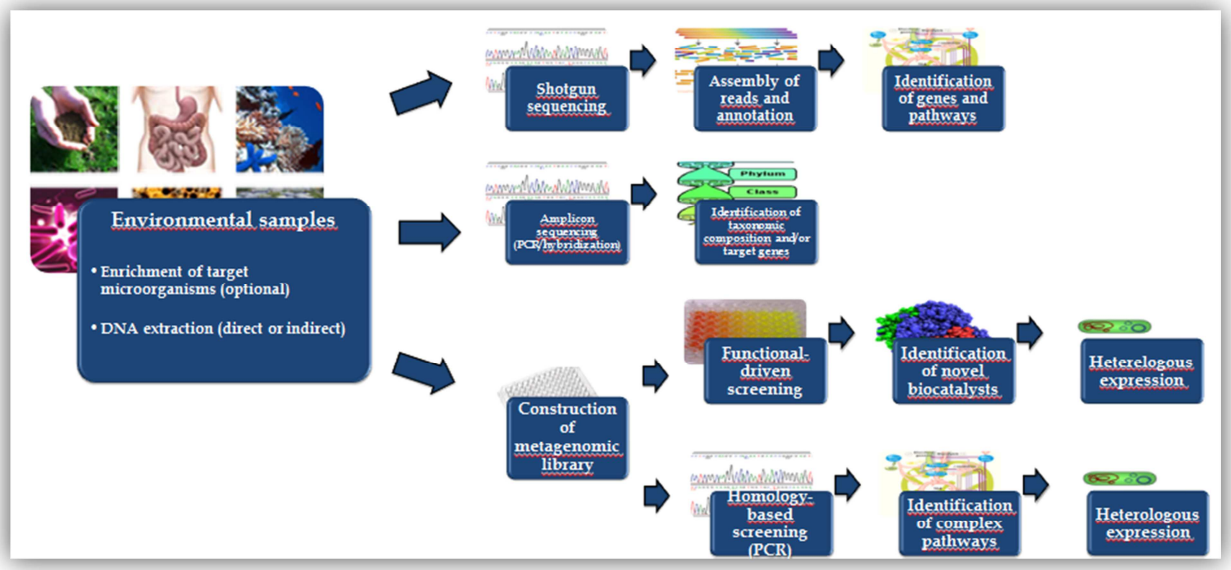


Figure 10 - General scheme of metagenomic approaches.

Both screening methodologies have been shown to be useful for the discovery of novel biocatalysts for lignocellulose conversion, and they enabled the identification of several (hemi)cellulases and accessory enzymes involved in (hemi)cellulose hydrolysis.

The main recent achievements of metagenomic approaches aimed at the development of new biocatalysts for lignocellulose saccharification, the techniques adopted, their main bottlenecks, the new biocatalysts developed and their industrial potential are summarized in the review: **"Metagenomics for the development of**

new biocatalysts to advance lignocellulose saccharification for bioeconomic development" (Appendix II).

3.2.1 – Shotgun sequencing-driven analysis

The sequence-driven metagenomic approach was applied to discover novel biocatalysts for lignocellulose hydrolysis from the three dedicated lignocellulosic energy crops *Arundo donax*, *Eucalyptus camaldulensis* and *Populus nigra* after 135 of natural biodegradation in underwood (named T3ADSB, T3ESB and T3PSB, respectively). Since lignocellulosic materials are efficiently decomposed in natural environments by cooperation of autochthonous microbial communities (Feng et al., 2011), the ecosystems investigated in our study represent a reservoir of novel genes encoding enzymes involved in (hemi)cellulose degradation, necessary for the development of eco-compatible and economically favorable industrial process. The use of next-generation high-throughput sequencing technologies was applied to study the complex microbial community structure and taxonomic diversity of the analyzed biomasses and to investigate putative genes coding for enzymes potentially relevant to the degradation of plant biomass into fermentable sugars. This study provides high-quality results on the lignocellulose-adapted microbial communities and hydrolytic activity showing that the three investigated feedstock represent an important source of enzymes to enhance the conversion efficiency of lignocellulosic crops into fermentable sugars. Moreover, they are useful for the identification of not yet known genes suitable for industrial biotechnological application.

Results

- Data Statistics

The microbiota of three different lignocellulosic biomasses (*A. donax, E. camaldulensis* and *P. nigra*) were analysed by Illumina sequencing of the metagenomic DNAs obtaining a total of 11,208,388,400, 11,274,127,600 and 2,392,000 raw reads for *A. donax*, *E. camaldulensis* and *P. nigra*, respectively. Low-quality reads were excluded, and around 10.0 Gb, for *A. donax* and *E. camaldulensis* samples, and 2 Gb, for *P. nigra*, high quality and clean sequence reads were selected (Table 6).

The reads were assembled into 95,292, 159,184 and 33,805 contigs (cut-off value 500 bp) for *A. donax*, *E. camaldulensis* and *P. nigra* biomasses, respectively (Table 1). The N50 and N90 contig lengths ranged from 914 to 1,452 and from 546 to 583, respectively and the longest contig was 49,245, 650,642 and 85,030 in *A. donax*, *E. camaldulensis* and *P. nigra*, respectively (Table 6).

Table 6. Quality and statistical summary of sequencing and assembling.

Parameter	Biomass sample		
	A. donax	*E. camaldulensis*	*P. nigra*
Total reads (bp)	11,208,388,400	11,274,127,600	2,392,000
High quality reads (bp)	10,010,000,000 (89%)	10,010,000,000 (88%)	2,100,000 (88%)
Number of contigs	95,292	159,184	33,805
Length of contigs (bp)	111,530,551	143,424,210	40,937,098
N50 contig length (bp)	1,326	914	1,452

N90 contig length (bp)		574	546	583
Largest contig (bp)		49,245	650,642	85,030
Shortest contig (bp)		500	500	500
	PE	6,497,524	7,683,451	2,852,867
Mapping	SE	1,790,597	2,097,376	1,253,515
	Total (%)	14.77	17.45	33.14

- Microbial community composition of lignocellulosic biomasses

The high quality reads were aligned with the sequences to NR database at NCBI to investigate the microbial community composition. The three lignocellulosic biomass samples were shown dominated by bacterial species (approximately 90.0%), followed by eukaryota (4.62%, 7.03% and 0.92% in *A. donax*, *E. camaldulensis* and *P. nigra*, respectively) and unclassified sequences (4.39%, 3.09% and 4.68%, in *A. donax*, *E. camaldulensis* and *P. nigra*, respectively). A negligible percentage matched viral species in *A. donax* and *E. camaldulensis* (0.23% and 0.11%, respectively) (Table 7).

Table 7. Samples composition and relative abundance of each taxon at the super kingdom level mapping the high quality reads to the NT database (NCBI).

	Arundo donax	*Eucalyptus camaldulensis*	*Populus nigra*
Bacteria			
reads	2,701,793	3,173,452	2,232,969
total %	90.76%	89.77%	94.40%
Eukaryota			
reads	137,550	248,521	21,789
total %	4.62%	7.03%	0.92%
Archea			
reads	11	16	6
total %	0.00%	0.00%	0.00%
Viruses			
reads	6,777	3,819	0
total %	0.23%	0.11%	0.00%
Viroids			
reads	16	17	0
total %	0.00%	0.00%	0.00%
Not assigned			
reads	130,698	109,409	110,752
total %	4.39%	3.09%	4.68%

The relative abundances of microbial taxa were examined at the level of genera to determine the dominant taxa characterizing the composition of the bacterial communities of the different investigated plant species.

In total, sixteen different genera with an incidence ≥ 1 % were detected in the biomass materials, but only *Streptomyces*, *Pseudomonas*, *Agrobacterium*, *Xanthomonas* and *Stenotrophomonas* were detected in all samples (Fig. 11). In particular, the composition of microbial community in all biomass samples was strongly dominated by *Streptomyces* being the most abundant taxa recovered, accounting for 34.99%, 47.70% and 50.08% of the total microbial diversity in *A. donax*, *E. camaldulensis* and *P. nigra*, respectively. The other genera (*Pseudomonas*, *Agrobacterium*, *Xanthomonas* and *Stenotrophomonas*), retrieved in all lignocellulosic biomass samples, belong to *Proteobacteria* phylum. The relative abundance of these taxa was very variable showing a percentage ranging approximately from 1% to 7%, depending on lignocellulosic plant species (Fig. 11). The occurrence of the other bacterial taxa recovered in the lignocellulosic biomass was very low showing an abundance of approximately ≤1%, with the exception of *Pantoea* (3.61% in *A. donax*), *Rahnella* (5.25% and 3.32% in *A. donax* and *P. nigra*, respectively), *Erwinia* (6.32% in *E. camaldulensis*) and *Bacillus* (7.79% in *P. nigra*) (Fig. 11).

The relative abundances of fungal taxa accounted for 2.18%, 4.87% and 0.22% of the total biodiversity in *A. donax*, *E. camaldulensis* and *P. nigra*, respectively (Fig. 11). In detail, the incidence of all fungal genera identified in *A. donax* and *P. nigra* biomass was <1 %; while in *E. camaldulensis* biomass, *Penicillium* strongly dominated the eukaryotic biodiversity showing a relative abundance of 3.22% (Fig. 11).

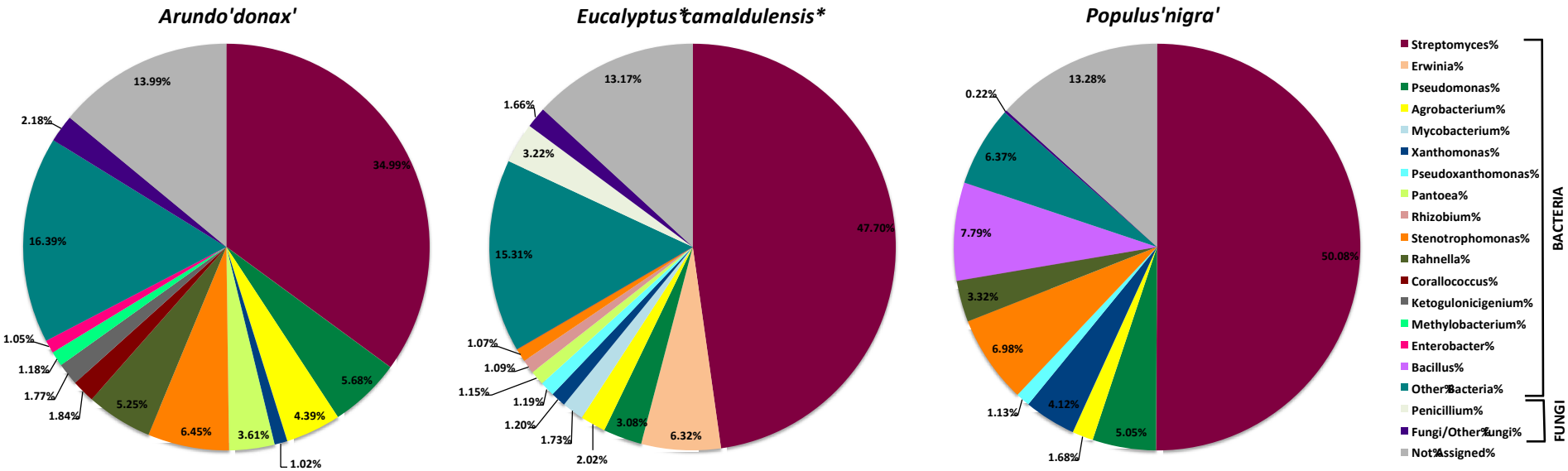


Figure 11. Abundance of bacterial and fungal genera in *A. donax, E. camaldulensis* and *P. nigra*. Only taxa with an incidence ≥1% in each sample are shown. Other bacteria and other fungi represent the aggregate of other bacterial and fungal genera, respectively; not assigned means that these reads cannot be annotated at the genus level.

- Microbial diversity related to Glycoside Hydrolases

Microbial diversity related to GH families of predicted ORFs from the three lignocellulosic biomasses was also investigated to identify the bacterial and fungal genera encoding enzymes involved in the carbohydrates metabolism. The microbial biodiversity related to GHs was very high and twenty-six bacterial and forty-two fungal genera were recovered with an incidence ≥ 1% in at least one sample (Fig. 12). *Streptomyces* was a dominant genus in all samples accounting for 18.10%, 28.28% and 30.02% in *A. donax*, *E. camaldulensis* and *P. nigra*, respectively, of the microbial genera related to GH (Fig. 12a).

Unlike the other biomasses in which *Streptomyces* was the dominant taxon, in *P. nigra* the most of GHs were related to *Paenibacillus* (30.38%) (Fig. 12a). *Pseudomonas* and *Rhizobium* were the other genera recovered in all lignocellulosic biomasses in relationship to the GHs (Fig. 12a) showing an abundance ranging from 1.08% to 6.30% depending on plant species.

The abundance of the other taxa related to GHs is strictly correlated to lignocellulosic plant species. In particular, a high percentage of bacteria belonging to *Stenotrophomonas* genus encoded GHs in *A. donax* (8.82%) and *P. nigra* (10.31%) (Fig. 12a). Moreover, the most of GHs in *A. donax* biomass was encoded by genera belonging to the class of *Actinobacteria* and in particular, *Curtobacterium* (5.96%), *Microbacterium* (8.71%), *Nocardiopsis* (5.96%) and *Promicromonospora* (1.15%) (Fig. 12a). Differently, members belonging to α-*Proteobcteria* (*Novosphingobium* and *Isoptericola*) and γ-*Proteobacteria* (*Pseudoxanthomonas*, *Xanthomonas*, *Dyella* and *Rhodanobacter*) classes characterized *E. camaldulensis* biomass; while γ-*Proteobacteria* (*Stenotrophomonas* and *Xanthomonas*) together to *Bacillus* (5.79%) were the other taxa recovered in the *P. nigra* biomass (Fig. 12a).

In this study, the GHs were related also to a wide range of fungal taxa. Among the forty-two genera occurring with abundance ≥ 1% in at least one sample, only *Pestalotiopsis* was recovered in all lignocellulosic biomasses (with an incidence of 2.70%, 2.06% and 12.50% in *A. donax*, *E. camaldulensis* and *P. nigra*, respectively) (Fig. 12b).

Overall, the highest fungal biodiversity related to GHs was found in *E. camaldulensis* (35 genera) followed by *A. donax* (13 genera) and *P. nigra* (5 genera). Although in *E. camaldulensis* the highest biodiversity was found, all fungal genera occurred at low percentage with the exception of *Nectria* (10.31%) and *Sporothrix* (20.62%) (Fig. 12b). By contrast, in *A. donax* biomass, the most of GHs were related to *Fusarium* (18.60%), *Nectria* (14.20%) and *Trichoderma* (12.40%), while the abundance of the other taxa range approximately from 8% to 1% (Fig. 12b).

Finally, the lowest fungal diversity was found in *P. nigra* biomass. The most abundant taxa recovered in this plant sample was *Togninia* (37.50%) followed by *Batrachochytrium* (25.50%) and *Pestalotiopsis*, *Meyerozyma* and *Ustilago* (12.50%) (Fig. 12b). However, although this result seemed to evidence a high incidence of these taxa, overall very few GHs were related to them because only the 0.22% of total biodiversity was determined by fungi in this sample (Fig. 11).

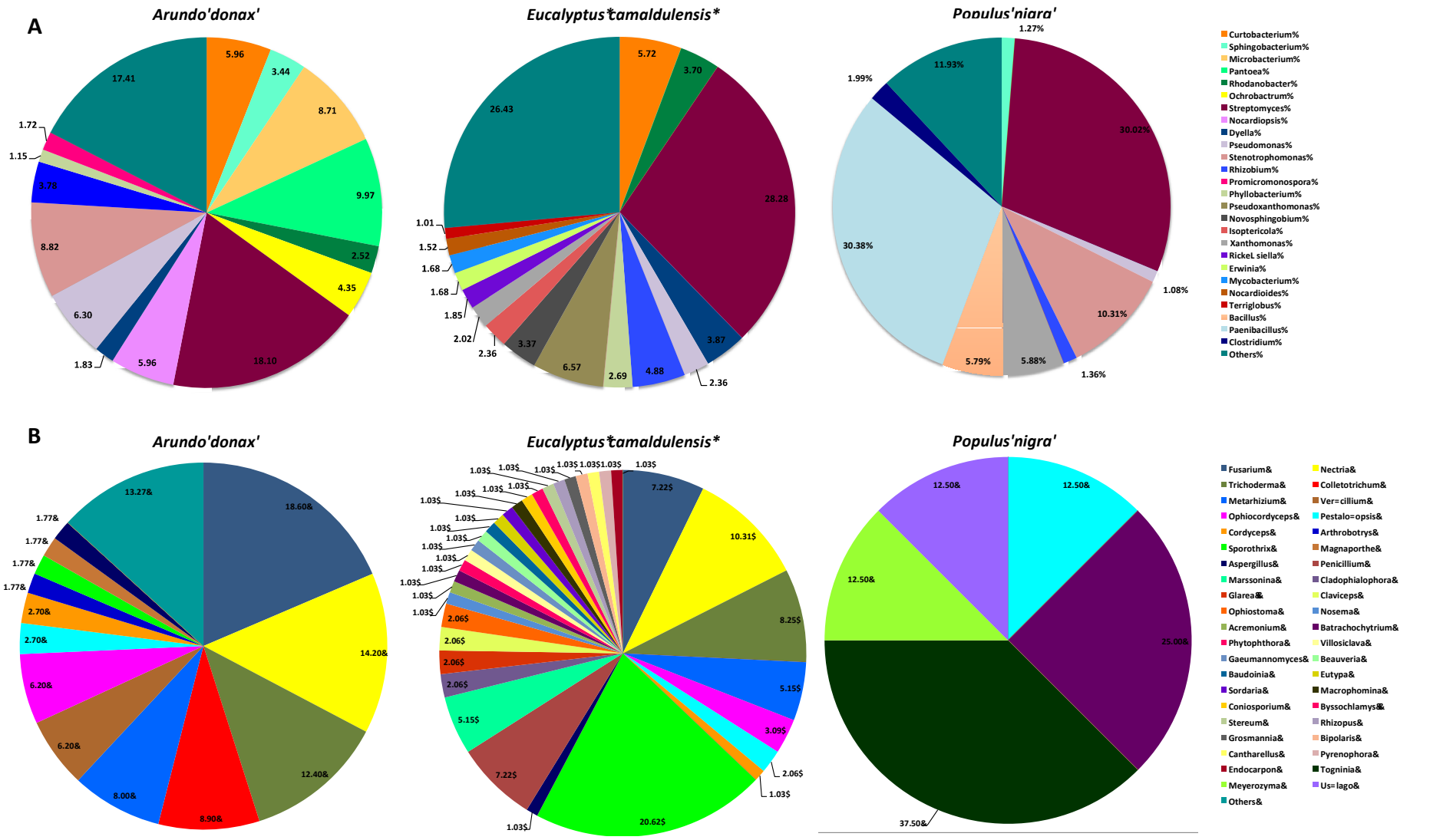


Figure 12. Percentage composition of bacterial (A) and fungal (B) genera related to GH families of predicted ORFs in *A. donax*, *E. camaldulensis* and *P. nigra*. Only taxa with an incidence ≥1% in each sample are shown. Others represent the aggregate of other bacterial (A) and fungal (B) genera.

- eggNOG and KEGG functional profiling of lignocellulosic biomass

With the aim to investigate the functional diversity in the three samples, the predicted amino acid sequences were also aligned to the databases Evolutionary genealogy of Genes non-supervised orthologous groups – eggNOG – and Kyoto Encyclopedia of Genes and Genomes – KEGG – by using BLAST.

The eggNOG database contains automatically annotated non-supervised orthologous groups with functional descriptions derived by identifying common denominators for the genes based on their various annotations (Jensen et al., 2008; Powell et al., 2013). This classification provides information about the functional annotation of orthologous genes with conserved function in different species. As shown in figure 13, the three samples are very similar with respect to the statistic results of the predicted ORFs alignment to eggNOG database. The data revealed a prevalence of poorly characterized genes belonging to S (function unknown) or R (general function prediction only) eggNOG category. Moreover, for all three samples a high percentage (~38%) of genes matching to non-supervised orthologous groups were classified involving in metabolism (categories C, E, F, G, H, I, P, Q) with ~8% of genes related to the carbohydrates transport and metabolism.

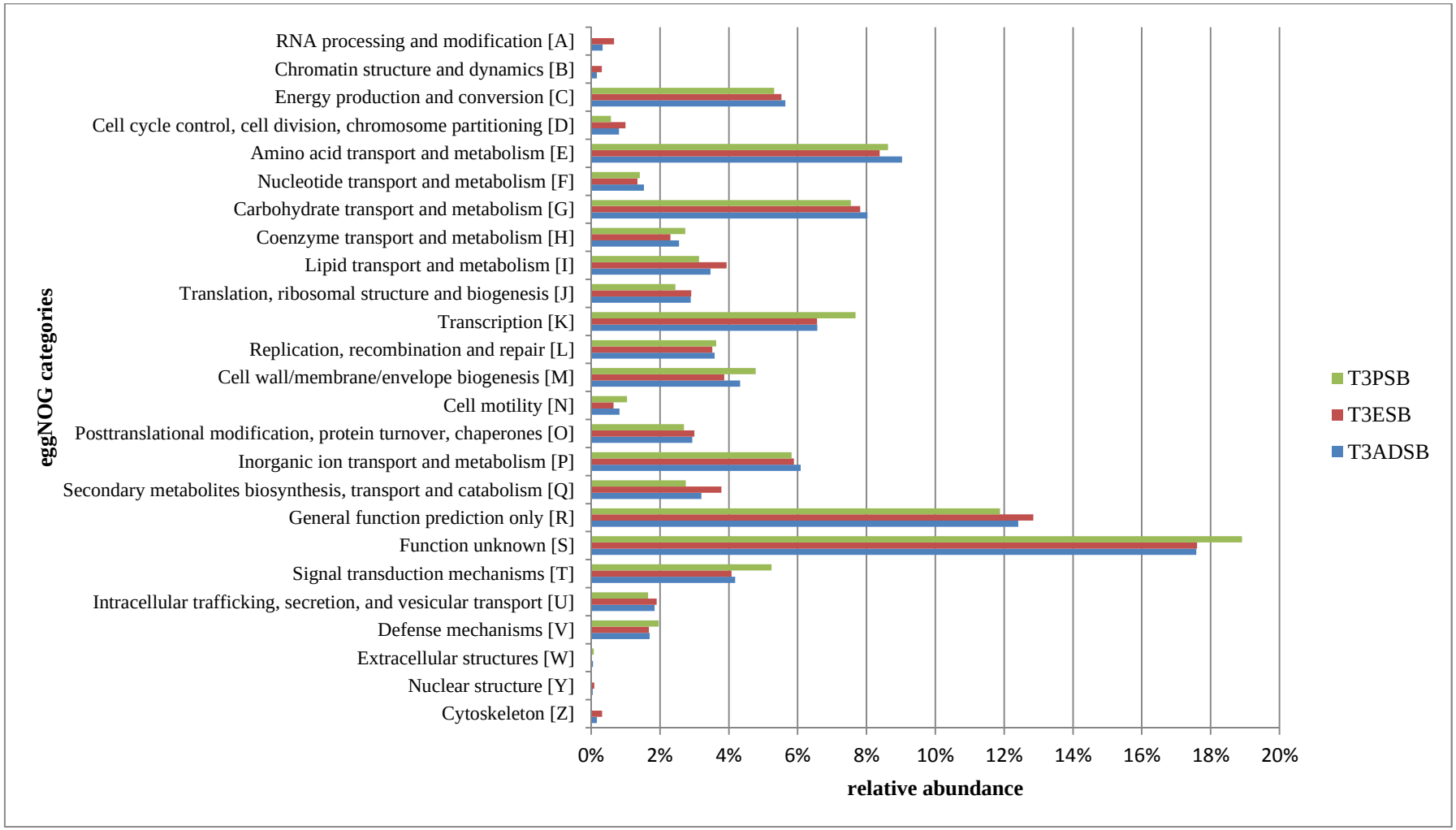


Figure 13. Relative abundance of eggNOG categories related to the predicted ORFs from T3ADSB, T3ESB and T3PSB sample.

The KEGG database is a collection of information from fully sequenced genomes, genes, proteins, pathways and chemical compounds concerning over a hundred different species. As shown in figure 14, even the KEGG classification profiles are very similar for all three samples and, although the majority of predicted ORFs were related to the membrane transport, this analysis confirms that a high abundance of genes matching to KEGG database (~12%) originated from pathways involved in the carbohydrates metabolism.

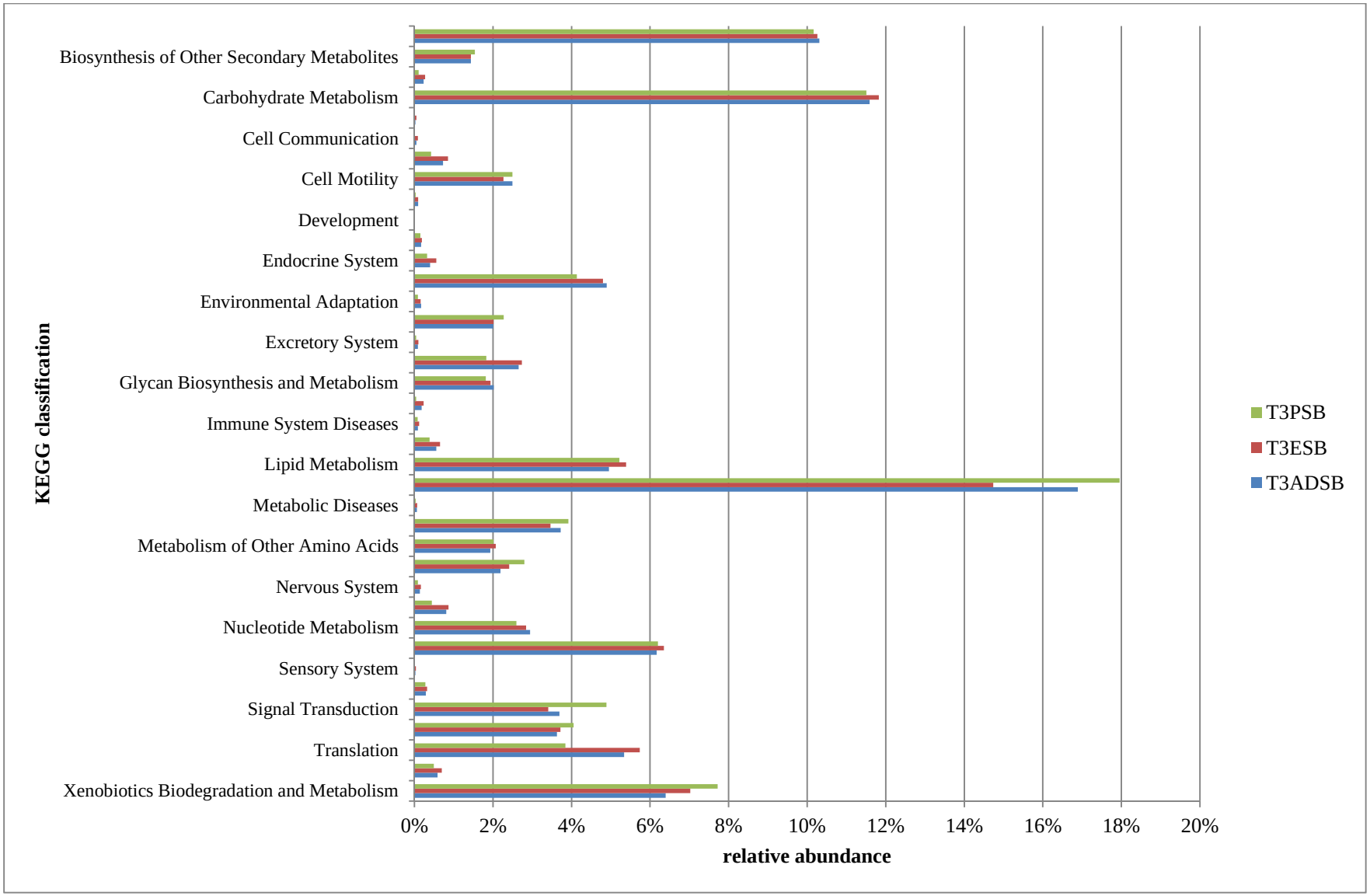


Figure 14. KEGG pathway classification of the predicted ORFs from T3ADSB, T3ESB and T3PSB samples.

- Inventory of the detected Carbohydrate-Active Enzymes families and putative plant-polysaccharides-targeting Glycoside Hydrolases

In order to identify putative genes and enzymes involved in breakdown, biosynthesis or modification of carbohydrates, the total predicted ORFs in the three investigated biomass samples were blasted against Carbohydrate-Active Enzymes (CAZymes) database. A total of 1792, 1279 and 2113 putative CAZymes were identified in the samples T3ADSB, T3ESB and T3PSB respectively, corresponding to 1.15%, 0.59% and 3.44% of the total ORFs (Table 8).

Table 8. CAZYmes classification of predicted ORFs from T3ADSB , T3ESB and T3PSB sample.
**the total numbers of CAZYmes is less than the sum (AAs+CBMs+CEs+GHs+GTs+PLs) due to the fact that some multimodular predicted proteins were detected.*

CAZymes classification	T3ADSB		T3ESB		T3PSB	
	# ORFs	%	# ORFs	%	# ORFs	%
Auxiliary Activities enzymes (AAs)	76	4,2%	23	1,8%	45	2,1%
Carbohydrate-Binding Modules (CBMs)	94	5,2%	148	11,6%	285	13,5%
Carbohydrate Esterases (CEs)	110	6,1%	68	5,3%	159	7,5%
Glycoside Hydrolases (GHs)	1059	59,1%	750	58,6%	1136	53,8%
Glycosyl Transferases (GTs)	460	25,7%	320	25,0%	555	26,3%
Polysaccharide Lyases (PLs)	24	1,3%	37	2,9%	48	2,3%
Total CAZYmes*	1792*		1279*		2113*	

A high relative abundance (25-26%) of predicted CAZymes was reported belonging to Glycosyl Transferases (GTs) families and involved in forming glycosidic bonds for the biosynthesis of di-, oligo- and polysaccharides. A less amount of Carbohydrate Esterases (CEs), Polysaccharide Lyases (PLs) and Auxiliary Activities (AAs) enzymes were detected in the three samples. Moreover, ORFs coding for putative Carbohydrate-binding modules (CBMs) were 5.2%, 11.6% and 13.5% on total CAZYmes for T3ADSB, T3ESB and T3PSB, respectively. CBMs are non-catalytic proteins contiguous within a CAZYme having binding activity to carbohydrates and helping the enzyme to degrade refractory substrates. The CBMs modules are classified into A, B or C types based on the nature of target substrates (crystalline polysaccharides, soluble polysaccharides or soluble oligosaccharides, respectively). Around half of the detected CBMs (2.5%, 6.3% and 7% on total CAZYmes for T3ADSB, T3ESB and T3PSB, respectively) was in conjunction with other non-catalytic CBMs and/or with catalytically-active GHs modules exhibiting a modular structure. In particular, in the metagenome from *A. donax*, most of the multimodular CAZYmes contained two modules, mainly a CBM48 module with glycogen-binding function at the N-terminal side of the GH13 module or multiple CBM modules from the same family 50, also known as LysM domains. Only 4 ORFs codifying for putative multimodular proteins contained three modules. In the metagenome from *E. camaldulensis*, multimodular CAZYmes containing CBM32 module were mainly detected. The members belonging to CBM family 32, commonly found in bacterial CAZymes that modify plant cell wall polysaccharides and eukaryotic glycans, are reported to have different substrate specificity (Abbott et al., 2008). Modular proteins containing CBM32 module were mainly detected even in metagenome from *P. nigra* in multiple copies within the same enzyme or in conjunction with other CBM and/or GH motifs. In this sample, the largest amount of multimodular CAZYmes was recognized. In particular, one ORF

consisted of seven modules: one GH16 and four CBM4 (reported acting on xylan, glucans and amorphous cellulose), followed by two CBM32 at the C-terminus of Glycoside Hydrolase – one of six modules – one CBM54 (one of the most recently discovered carbohydrate-binding modules) at the N-terminal side of a GH16 and four CBM4 at the C-terminus – and one of five modules – four CBM35 at the C-terminal side of a GH43.

However, most of the detected CAZymes in the three samples were involved in hydrolysis and/or rearrangement of glycosidic bonds. In particular, a number of 1059 in *A. donax* (corresponding to 59.10% on total CAZymes and to 0.68% on total ORFs detected), 750 in *Eucalyptus camaldulensis* (corresponding to 58.60% on total CAZymes and to 0.34% on total ORFs detected) and 1136 in *Populus nigra* (corresponding to 53.80% on total CAZymes and to 1.85% on total ORFs detected) predicted proteins were classified as Glycoside Hydrolases (GHs). The figures 15, 16 and 17 show the most frequently occurring putative GHs detected in T3ADSB, T3ESB and T3PSB samples, respectively. For each sample, the GHs with more than 1% of abundance of the total detected GHs are reported.

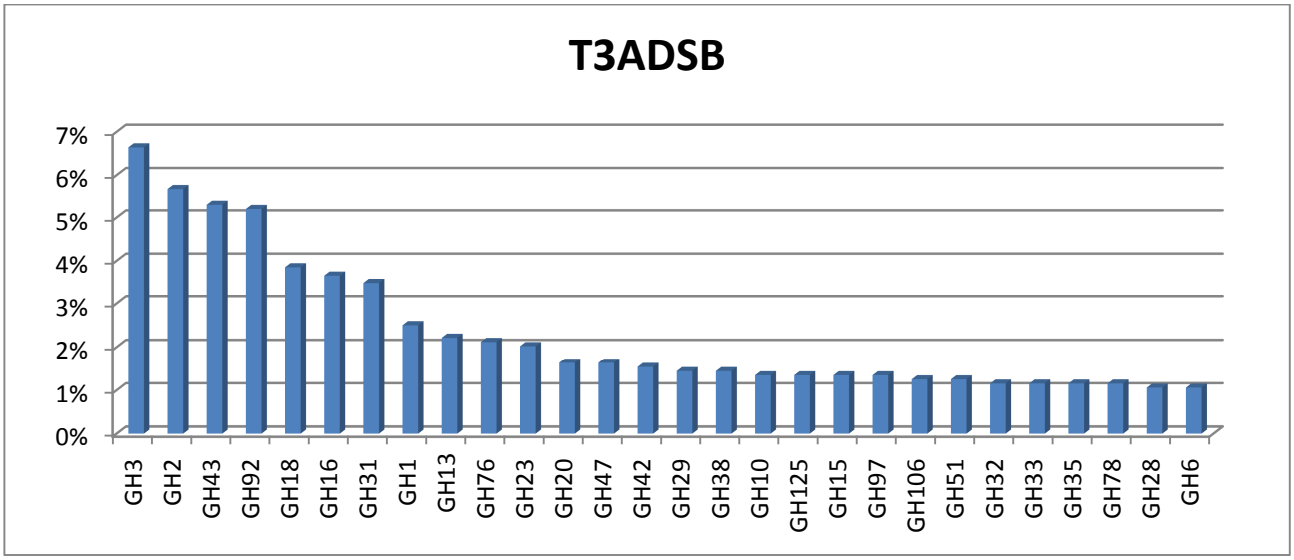


Figure 15. GH family percentage of predicted ORFs from T3ADSB sample. The GHs with more than 1 % of abundance are reported.

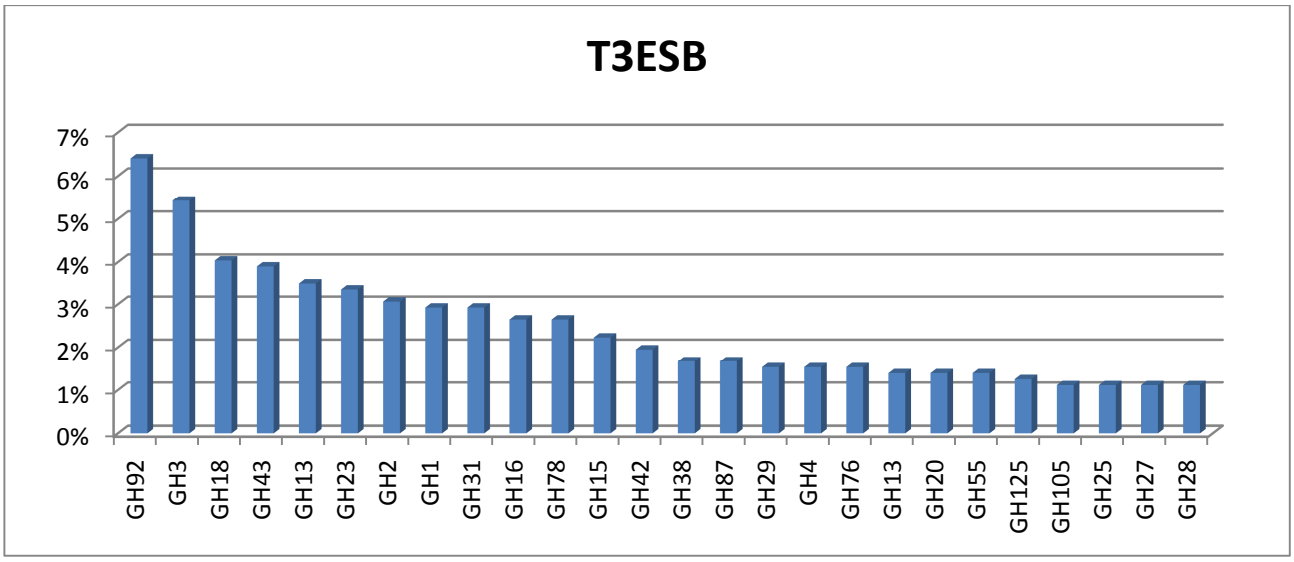


Figure 16. GH family percentage of predicted ORFs from T3ESB sample. The GHs with more than 1 % of abundance are reported.

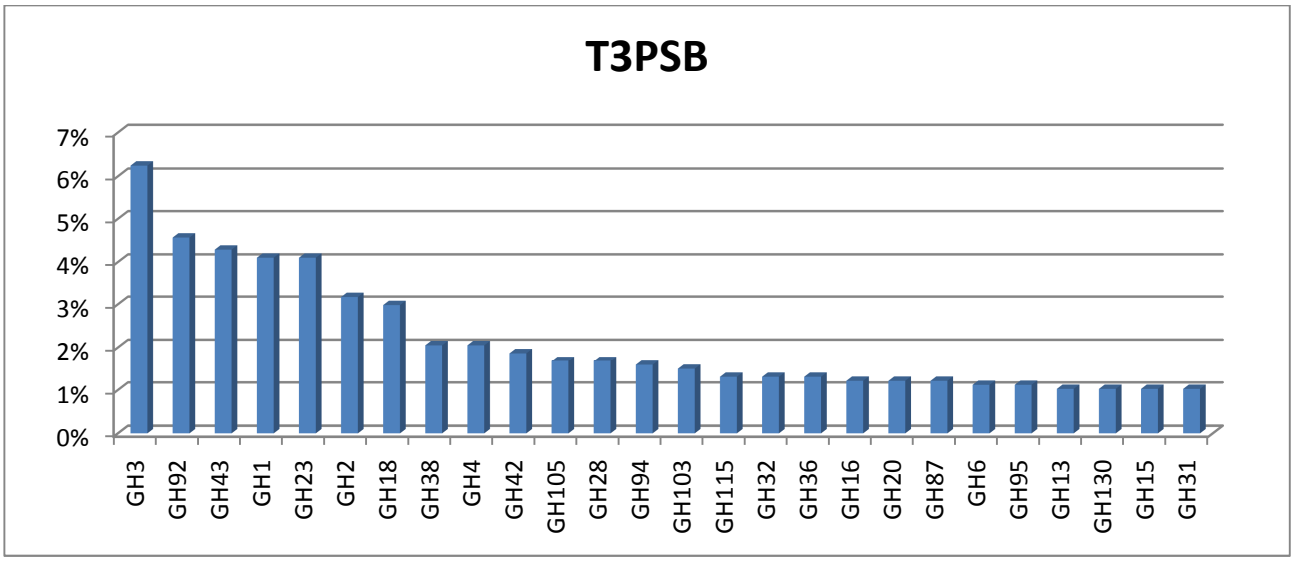


Figure 17. GH family percentage of predicted ORFs from T3PSB sample. The GHs with more than 1 % of abundance are reported.

Table 9 shows the comparison of GH family percentage (with more than 3% of abundance of the total) of predicted ORFs from the samples. An abundance of putative GH92 (exo-acting α-mannosidases), GH3 and GH43 is noted in all samples. Moreover, in the sample from *Arundo donax* and *Eucalyptus camaldulensis*, a large amount of GH18 was detected. This family is reported to include both chitinases and endo-β-N-acetylglucosaminidases but also sub-families of non-hydrolytic proteins. In the metagenome from *Eucalyptus camaldulensis*, a relative abundance of CAZYmes belonging to GH13 family was detected. The GH13 enzymes are reported having a wide range of different target substrates and products and have been subdivided into almost 40 subfamilies, most of which are monofunctional. In particular, in our samples , only GH13 belonging to subfamily 11 (reported having debranching activity on glycogen, amylopectin and their β-limit dextrins) and subfamily 30 (involved in the hydrolysis of terminal α-D-glucose residues with release of monomers) were detected. The sample from *Populus*

nigra shows even a high abundance of GHs belonging to GH23 family. All the enzymes belonging to GH23 family are reported to have activity on peptidoglycan and, in particular, the lysozymes to have activity even on chitin and chitooligosaccharides.

Table 9. Comparison of GH family percentage of predicted ORFs from T3ADSB , T3ESB and T3PSB sample. The GHs with more than 3 % of abundance are reported.

CAZy Family	Main Known activities	T3ADSB	T3ESB	T3PSB
GH1	β-glucosidases, β-galactosidases, 6-phospho-β-glucosidase and 6-phospho-β-galactosidase, β-mannosidase, β-D-fucosidase and β-glucuronidase	2.50%	2.92%	4.10%
GH2	β-galactosidases, β-glucuronidases, β-mannosidases, exo-β-glucosaminidases	5.68%	3.06%	3.17%
GH3	exo-acting β-D-glucosidases, α-L-arabinofuranosidases, β-D-xylopyranosidases and N-acetyl-β-D-glucosaminidases	6.64%	5.43%	6.24%
GH16	Xyloglucosyltransferase, keratan-sulfate endo-1,4-β-galactosidase, endo-1,3-β-glucanase, endo-1,3(4)-β-glucanase, licheninase, β-agarase, κ-carrageenase, xyloglucanase, endo-β-1,3-galactanase, β-porphyranase, hyaluronidase, endo-β-1,4-galactosidase, chitin β-1,6-glucanosyltransferase, endo-β-1,4-galactosidase	3.66%	2.65%	1.21%
GH18	chitinases and endo-β-N-acetylglucosaminidases	3.85%	4.04%	2.98%
GH23	muramidase, peptidoglycan N-acetylmuramoylhydrolase, 1,4-β-N-acetylmuramidase and N-acetylmuramoylhydrolase	2.02%	3.34%	4.10%
GH43	α-L-arabinofuranosidases, endo-α-L-arabinanases (or endo-processive arabinanases) and β-D-xylosidases	5.29%	3.90%	4.29%
GH92	exo-acting α-mannosidases	5.20%	6.41%	4.57%

- KEGG pathway classification related to Glycoside Hydrolases

An in-depth KEGG pathway mapping was carried out for the putative genes coding for Enzyme Commission (EC) number activities related to the hydrolysis of glycosidic. As shown in figure 18, a high percentage of different cellulases were detected. In particular, β-glucosidases (EC 3.2.1.21, hydrolyzing cellobiose and other cellodextrins) and endo-1,4-β-glucanases (EC 3.2.1.4, performing the random internal hydrolysis of amorphous cellulose) were the most abundant putative enzymes involved in the hydrolysis of glycosidic bonds. In the samples from *Arundo donax* and *Populus nigra,* a relative abundance of chitinases (EC 3.2.1.14) was also detected (7.41% and 6.29% respectively). These enzymes, involved in the degradation of chitin - a linear biopolymer of β-1, 4-N-acetylglucosamine (GlcNAC), abundantly present in fungi, insects and crustaceans - are even produced by several soil-borne microorganisms to use chitin as a source of energy, in order to supply nitrogen and carbon (Hamid et al., 2013). It is noteworthy that in all three samples, putative genes coding for hemicellulases and accessory enzymes with a broad spectrum of activities were recognized. In particular, a significant percentage of proteins involved in the degradation of (glucurono)(arabino)xylan - such as endoxylanases (EC 3.2.1.8) and β-xylosidases – and in the removal of arabinose - α-L-arabinofuranosidases (EC 3.2.1.55) - or galactose - α-galactosidases (EC 3.2.1.22) - substituents in hemicelluloses were detected. Moreover, several additional putative enzymes related to the hemicelluloses degradation – such as mannanases (EC 3.2.1.78), polygalacturonases (EC 3.2.1.67) and feruloyl esterases (EC 3.1.1.73) were recognized in a lower percentage.

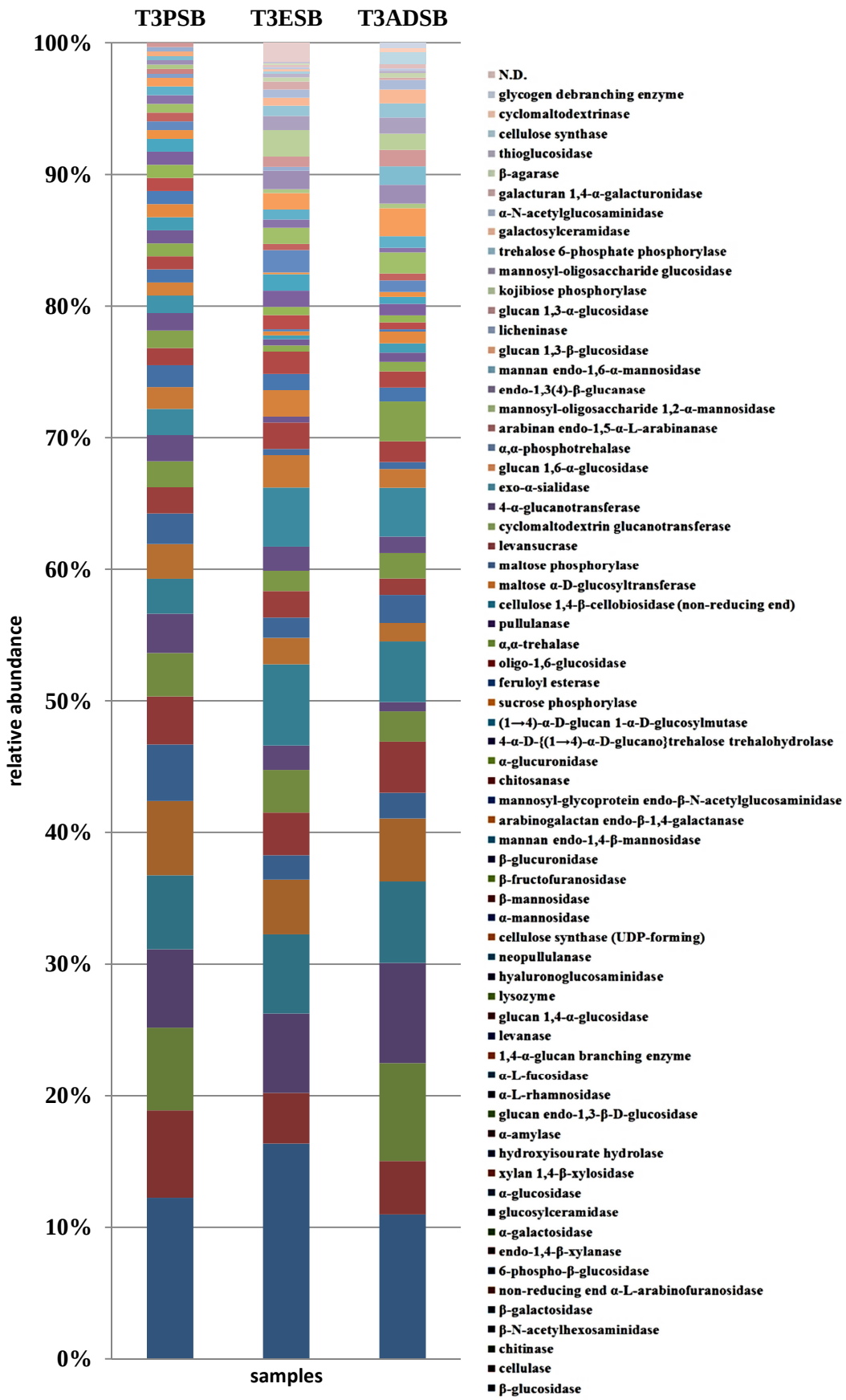


Figure 18. KEGG pathway classification for the putative genes coding for Enzyme Commission (EC) number activities related to the hydrolysis of glycosidic bonds from T3ADSB, T3ESB and T3PSB samples.

Discussion

In the last decades, the increasing interest in the use of renewable sources for green energy and chemicals has strongly stimulated search for new biocatalysts from different ecosystems for lignocellulose conversion. Therefore, microbial and enzymatic diversities potentially relevant to the degradation of plant biomass into fermentable sugars were explored through metagenomic approach in three dedicated lignocellulosic energy crops, *Arundo donax*, *Eucalyptus camaldulensis* and *Populus nigra*, after natural biodegradation. Microbial DNA sequences were analysed to assess the total biodiversity, identify putative genes and enzymes involved in carbohydrates metabolism and evaluate bacterial and fungal diversity related to GH families of predicted ORFs

The microbiomes of the three lignocellulosic biomasses were shown to be strongly dominated by bacterial taxa, in which the most frequently occurring bacteria were those belonging to the *Streptomyces* genus. Its predominance could be due to the ability to synthetize enzymes, such as endocellulases, laccases and peroxidases (Amore et al., 2012; Woo et al., 2014), which efficiently degrade lignocellulosic materials under a wide range of environmental conditions (Wang et al., 2014). Recently, *Actinobacteria*, and in particular *Streptomyces* spp. are also reported among the most efficient plant biomass-degrading microbes in peat swamp forests as well as the dominant class during the composting process of chestnut green waste (Kanokratana et al., 2011; Ventorino et al., 2016). Bacterial species belonging to *Proteobacteria* phylum, such as *Pseudomonas* spp. and *Stenotrophomonas* spp., were also retrieved in all lignocellulosic samples. Bacteria belonging to these genera are able to produce a wide range of enzymes for efficient degradation of carboxymethylcellulose, (hemi)cellulose and lignin (Talia et al., 2012; Wang et al., 2013). Moreover, *Actinobacteria* and *Proteobacteria* have been identified as the predominant bacterial phyla during composting of lignocellulosic waste exhibiting the enzymatic activities required for the degradation of this recalcitrant polymeric material (Lopez-Gonzalez et al., 2015).

The occurrence of other bacterial taxa was also demonstrated in the investigated lignocellulosic biomasses with a different abundance depending on plant species. Interestingly, *Bacillus* genus covered approximately 8.0% of the total microbial biodiversity in *P. nigra*. Members belonging to *Bacillus* spp. isolated from different environments exhibit cellulolytic and/or hemicellulolytic activities as well as multi-enzyme complexes to breakdown the components of lignocellulosic material (Jones et al., 2012; Amore et al., 2013a, 2013b; Di Pasqua et al., 2014). Moreover, different microbial strains belonging to *Enterobacteriaceae* family such as *Pantoea*, *Rahnella* and *Erwinia*, are frequently recovered in the gut of insects producing digestive enzymes implicated in the hydrolysis of cellulose (Anand et al., 2010; Morales-Jiménez et al., 2012).

Moreover, an interesting finding was the low abundance of eukaryotic populations in the total biodiversity of the lignocellulosic biomass samples. Among all fungal taxa retrieved, only *Penicillium* showed an incidence > 1% in *E. camaldulensis*. Cellulolytic activity of this genus is well documented and there are several reports on β-glucosidase, cellulases and xylanases production from different *Penicillium* species (Jorgensen et al., 2003; Camassola et al., 2008; Krogh et al., 2010). Moreover, Ryckeboer et al. (2003) reported also the ability of *Penicillium* spp. to degrade lignin and starch making it a good candidate in the producing of industrial cellulases (Adsul et al., 2011).

Analysing the biodiversity related to GH families of predicted ORFs, a highly complex microbial community was found. With regard to bacterial biodiversity, *Streptomyces*, *Pseudomonas* and *Rhizobium* were found in all lignocellulosic biomass samples. In agreement with the results obtained analysing the total biodiversity, *Streptomyces* was the

dominant taxon, confirming the ability of the members belonging to this genus to encode enzymes involved in cellulose and hemicellulose degradation. In fact, *Streptomyces* spp. is known able to produce different GHs that are well characterized (Goedegebuur et al., 2002; Ducros et al., 2000). In addition, the production of cellulolytic enzymes in *Rhizobium* spp. is related to their ability to nodulate leguminous plants. In fact, *Rhizobium* is a plant growth promoting rhizobacterium living as free-living saprophytes in the soil but also able to fix nitrogen establishing a symbiotic associations with a host plant (Ventorino et al., 2012). Therefore, the production of enzymes, such as cellulases, is fundamental to degrade plant cell wall polymers and penetrate in the host root (Robledo et al., 2008). García-Fraile et al. (2007) reported the ability to actively hydrolyse CMC of two bacterial strains isolated from decaying wood of *Populus alba* and classified as *Rhizobium cellulosilyticum*.

However, the prokaryotic biodiversity related to GHs was also dominated by *Paenibacillus* genus in the *P. nigra* biomass. Recently, a new endoglucanase having catalytic domain belonging to the glycoside hydrolases family 9 was found in *Paenibacillus* sp. (Fu et al., 2010). Eida et al. (2012) reported the ability of different *Paenibacillus* isolates to efficiently contribute to cellulolytic and hemicellulolytic processes during composting of sawdust. Moreover, cellulolytic enzymes from *Paenibacillus* spp., have been also applied in pulp refining (Pastor et al., 2001). The other taxa recovered in the *P. nigra* biomass, such as *Stenotrophomonas* and *Xanthomonas* (*Proteobacteria*) and *Bacillus* (*Firmicutes*) are known as plant biomass-degrading microbes. De Angelis at al. (2010) reported that the members of *Proteobacteria* as well as *Firmicutes* strongly dominated switchgrass-adapted communities comprising approximately 80% of the microbial richness.

Differently, in *A. donax* biomass the most of GHs was encoded by genera belonging to the class of *Actinobacteria.* These taxa are related to well characterized potent plant polysaccharide-degrading bacteria and play an important role in degradation of numerous polymers such as chitin, cellulose, lignin and polyphenol (Castillo et al., 2013).

With regard to fungal biodiversity related to GHs, a high variety of genera were found, and among these only *Pestalotiopsis* was recovered in all lignocellulosic biomasses. This result is in agreement with Cahyani et al. (2004) that reported the ubiquitous presence of *Pestalotiopsis* spp. during the composting process of rice straw. In fact, this endophytic fungus is able to secrete xylanases and cellulases also in salt stress conditions (Arfi et al., 2013) as well as produce a considerable amount of ligninolytic enzymes such as laccase (Chen et al., 2011).

However, *Sporothrix*, *Fusarium*, *Nectria* and *Trichoderma* dominated the eukaryotic biodiversity related to GHs in *A. donax* and *E. camaldulensis* biomasses. These Ascomycota are known for their ability to produce cellulolytic enzymes (Gherbawy and Abdelzaher, 1999; Wenzel et al., 2002) and comprise many species involved in the degradation of recalcitrant substances such as cellulose, hemicellulose, pectin, and lignin (Li et al., 2013). Jurado et al. (2014) reported that fungi belonged to Ascomycota group were ubiquitous throughout the whole lignocellulose-based composting process.

The functional clustering of the predicted ORFs to eggNOG and KEGG databases showed high similarity among the three analyzed samples. In particular, the prevalence of poorly characterized genes obtained by matching to eggNOG categories suggest the three detected biomasses as potential sources of not yet known genes. Moreover, the linkage of the genes to the related gene products in a specific pathway by using KEGG database allowed to integrate the knowledge about the functional classification of the detected ORFs. The high relative abundance of genes linked to carbohydrates metabolism pathway was higher than or similar to that detected in metagenomes from samples with well-known

lignocellulose-degrading ability, such as invasive snail crop microbiome (Cardoso et al., 2012) and lower termite *Coptotermes gestroi* gut (Do et al., 2014). This result confirm the high potentiality of the three analyzed metagenomes to express genes involved in lignocellulosic biomasses biotransformation.

Moreover, the inventory of the Carbohydrate-Active Enzymes families detected in the three samples interestingly revealed ORFs codifying for putative lytic polysaccharide monooxygenases (LPMOs). Nowadays, the interest is moving towards the LPMOs belonging to AA9, AA11 (both formerly reported as GH61) or AA10 (formerly reported as CMB33) families, due to their ability to depolymerize the recalcitrant insoluble polysaccharides from highly crystalline cellulose, increasing the efficiency of lignocellulose saccharification (Beeson et al., 2015). Only a few of LPMOs have been discovered by metagenomic approach (Montella et al., 2015). In metagenomes analyzed in this study, 3, 5 and 9 ORFs (for T3ADSB, T3ESB and T3PSB respectively) were assigned to AA10 family, whereas only in the sequenced eDNA from *A. donax* 11 and 2 ORFs codifying for putative enzymes belonging respectively to AA9 and to AA11 families were detected.

However, most of CAZymes detected in the three samples were related to putative plant-polysaccharides-targeting GHs. Based on the results of 46 finished metagenomic studies collected in Genomes OnLine Database (GOLD) and reported by Li et al. (2009), the percentages of detected GHs are higher than those present in metagenomic samples from soil, sludge and marine or lake environments. Moreover, the GHs abundance in T3PSB sample was higher even than that detected in microbiota of different organisms – such as plant-fed elephant (Ilmbergeer et al., 2014), invasive snail (Cardoso et al., 2012), cow (Brulc et al., 2009), termite (Warnecke et al., 2007) – generally showing more predicted genes involved in hydrolysis of the glycosidic linkage than those from every other samples. Furthermore, the diversity of GH family enzymes detected in the three samples was higher than that observed in samples with the highest lignocellulose-degrading potentiality (Ilmbergeer et al., 2014), in line with the detected high phylogenetic diversity. Based on the identification by pfam HMMs and the grouping proposed by Allgaier et al. (2010) according to the main enzymatic activity, the putative genes encoding proteins involved in the degradation of plant polysaccharides were detected and compared to the environmental metagenomic samples that showed the highest lignocellulose-degrading potentiality (Table 10).

CAZY family	Main known activity	Pfam domain	Metagenomic samples																			
			T3ADSB		**T3ESB**		**T3PSB**		**Six years old elephant feces** (Ilmbergeer et al., 2014)		**Yak rumen** (Dai et al., 2012)		**Snail crop (**Cardoso et al., 2012)		**Cow rumen** (Hess et al., 2011)		**Bovine rumen** (Brulc et al., 2009)		**Macropod gut** (Pope et al., 2010)		**Termite hindgut** (Warnecke et al., 2007)	
Endo-Celluases																						
GH5	cellulase	PF00150	40	3,8%	15	2,0%	23	2,0%	517	4,7%	1302	3,5%	36	1,4%	1451	5,2%	7	0,7%	10	1,8%	56	8,0%
GH6	endoglucanase	PF01341	11	1,0%	5	0,7%	12	1,1%	0	0,0%	0	0,0%	4	0,2%	0	0,0%	0	0,0%	0	0,0%	0	0,0%
GH7	endoglucanase	PF00840	11	1,0%	1	0,1%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	1	0,0%	0	0,0%	0	0,0%	0	0,0%
GH9	endoglucanase	PF00759	2	0,2%	3	0,4%	5	0,4%	119	1,1%	767	2,0%	15	0,6%	795	2,9%	6	0,6%	0	0,0%	9	1,3%
GH44	endoglucanase	NA	0	0,0%	0	0,0%	0	0,0%	7	0,1%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	6	0,9%
GH45	endoglucanase	PF02015	1	0,1%	0	0,0%	0	0,0%	7	0,1%	13	0,0%	0	0,0%	115	0,4%	0	0,0%	0	0,0%	4	0,6%
GH48	endo-processive cellulase	PF02011	1	0,1%	1	0,1%	3	0,3%	0	0,0%	32	0,1%	2	0,1%	3	0,0%	0	0,0%	0	0,0%	0	0,0%
total			66	6,2%	25	3,3%	43	3,8%	650	5,9%	2114	5,6%	57	2,2%	2365	8,5%	13	1,4%	10	1,8%	75	10,7%
Endo-hemicellulases																						
GH8	endo-xylanases	PF02011	1	0,1%	1	0,1%	7	0,6%	85	0,8%	174	0,5%	46	1,8%	329	1,2%	4	0,4%	1	0,2%	5	0,7%
GH10	endo-1,4-β xylanase	PF00331	15	1,4%	5	0,7%	8	0,7%	258	2,3%	2664	7,1%	25	1,0%	1025	3,7%	7	0,7%	11	2,0%	46	6,5%
GH11	xylanase	PF00457	4	0,4%	2	0,3%	4	0,4%	20	0,2%	244	0,6%	1	0,0%	165	0,6%	1	0,1%	0	0,0%	14	2,0%
GH12	endoglucanase & xyloglucan hydrolases	PF01670	3	0,3%	1	0,1%	3	0,3%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%
GH26	β- mannanase & xylanase	PF02156	5	0,5%	6	0,8%	10	0,9%	103	0,9%	537	1,4%	11	0,4%	369	1,3%	5	0,5%	5	0,9%	15	2,1%
GH28	galacturonases	PF00295	11	1,0%	8	1,1%	20	1,8%	242	2,2%	244	0,6%	69	2,7%	472	1,7%	5	0,5%	2	0,4%	6	0,9%
GH53	endo-1,4-β-galactanase	PF07745	4	0,4%	1	0,1%	5	0,4%	88	0,8%	1066	2,8%	9	0,3%	0	0,0%	17	1,8%	9	1,6%	12	1,7%
total			43	4,1%	24	3,2%	57	5,0%	796	7,2%	4929	13,1%	161	6,2%	2360	8,5%	39	4,1%	28	5,0%	98	13,9%
Debranching enzymes																						
GH51	α-L-arabinofuranosidase	NA	14	1,3%	7	0,9%	13	1,1%	239	2,2%	0	0,0%	22	0,8%	0	0,0%	64	6,7%	12	2,2%	18	2,6%
GH54	α-L-arabinofuranosidase	PF09206	0	0,0%	3	0,4%	0	0,0%	13	0,1%	111	0,3%	0	0,0%	0	0,0%	1	0,1%	0	0,0%	0	0,0%
GH62	α-L-arabinofuranosidase	PF03664	3	0,3%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	2	0,1%	1	0,0%	0	0,0%	0	0,0%	0	0,0%
GH67	α-glucuronidase	PF07477 PF07488	0	0,0%	3	0,4%	8	0,7%	0	0,0%	1090	2,9%	5	0,2%	120	0,4%	0	0,0%	5	0,9%	10	1,4%
GH78	α-L-rhamnosidase	PF05592	0	0,0%	19	2,5%	7	0,6%	413	3,7%	426	1,1%	73	2,8%	1260	4,5%	34	3,6%	25	4,5%	0	0,0%

total			17	1,6%	32	4,3%	28	2,5%	665	6,0%	1627	4,3%	102	3,9%	1381	5,0%	99	10,3%	42	7,5%	28	4,0%
Oligosaccharide-degrading enzymes																						
GH1	β-glucosidase & other β-linked dimers	PF00232	26	2,5%	23	3,1%	46	4,0%	103	0,9%	331	0,9%	294	11,4%	253	0,9%	10	1,0%	61	11,0%	22	3,1%
GH2	β-galactosidases & other β-linked dimers	PF02836 PF00703 PF02837	59	5,6%	22	2,9%	37	3,3%	917	8,3%	942	2,5%	66	2,5%	1436	5,2%	186	19,4%	24	4,3%	23	3,3%
GH3	mainly βglucosidases	PF00933	69	6,5%	39	5,2%	70	6,2%	804	7,3%	5448	14,5%	219	8,5%	2844	10,2%	176	18,4%	72	12,9%	69	9,8%
GH29	α-L-fucosidase	PF01120	15	1,4%	11	1,5%	9	0,8%	376	3,4%	899	2,4%	70	2,7%	939	3,4%	74	7,7%	2	0,4%	0	0,0%
GH35	β-galactosidase	PF01301	12	1,1%	7	0,9%	9	0,8%	123	1,1%	468	1,2%	32	1,2%	158	0,6%	12	1,3%	3	0,5%	3	0,4%
GH38	α-mannosidase	PF01074 PF07748	15	1,4%	12	1,6%	22	1,9%	81	0,7%	90	0,2%	18	0,7%	272	1,0%	17	1,8%	3	0,5%	11	1,6%
GH39	β-xylosidase	PF01229	7	0,7%	6	0,8%	4	0,4%	89	0,8%	159	0,4%	6	0,2%	315	1,1%	2	0,2%	1	0,2%	3	0,4%
GH42	β-galactosidase	PF02449 PF08533 PF08532	16	1,5%	14	1,9%	24	2,1%	37	0,3%	207	0,6%	54	2,1%	374	1,3%	11	1,1%	8	1,4%	24	3,4%
GH43	arabinases & xylosidases	PF04616	56	5,3%	28	3,7%	52	4,6%	894	8,1%	2313	6,2%	185	7,1%	0	0,0%	61	6,4%	10	1,8%	16	2,3%
GH52	β-xylosidase	PF03512	0	0,0%	0	0,0%	2	0,2%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	3	0,4%
total			275	26,0%	162	21,6%	275	24,2%	3424	31,0%	10857	28,9%	944	36,4%	6591	23,7%	549	57,4%	184	33,0%	174	24,7%
total plant plysaccharides targeting GHs			401	37,9%	243	32,4%	403	35,5%	5535	50,1%	19527	52,0%	1264	48,8%	12697	45,7%	700	73,1%	264	47,4%	375	53,3%
total GHs			1059		750		1136		11038		37563		2590		27755		957		557		704	

Table 10. Comparison of plant polysaccharides hydrolyzing enzymes between T3ADSB, T3ESB and T3PSB and samples with the highest lignocellulose-degrading potentiality.

The putative ORFs encoding enzymes related to the oligosaccharides degradation represented the majority of the total plant-polysaccharides-targeting GHs and their abundance (~26% for T3ADSB, ~22% for T3ESB and ~24% for T3PSB) was comparable to that detected in samples from cow rumen (Hess et al., 2011) and termite hindgut (Warnacke et al., 2007). Most belonged to GH1, GH2 and GH3 families including β-glucosidases, β-galactosidases, β-mannosidase, β-glucuronidase, β-xylosidase and other enzymes involved in the breakdown of a large variety of β-linked disaccharides. Due to the high diversity of protein structural arrangements, a robust phylogenetic classification of these families is currently not available. In addition, enzymes belonging to GH43 family were much represented (mainly in T3ADSB and T3PSB). This family includes β-xylosidases and α-L-arabinofuranosidases and several bifunctional enzymes; moreover, due to a remarkable expansion in GH43 family resulting from novel studies about plant cell wall degrading organisms, members of this family may have a more extensive range of specificities.

In the sample T3ADSB, the abundance of endocellulases was double than T3ESB and T3PSB and comparable to that detected in the six-years-old elephant feces (Ilmbergeer et al., 2014) and in yak rumen (Doi et al., 2012). The GH5 and GH6 were the most represented families. While only endoglucanase and cellobiohydrolase activities have been reported for the members of GH6 family, the enzymes belonging to Glycoside Hydrolases family 5 have a variety of specificities: this is one of the largest of all CAZy glycoside hydrolase families comprising not only cellulases, such as endo- and exo-glucanases and β-glucosidases, but even hemicellulases, such as endo- and exo-mannanases and β-mannosidase. Interestingly, in T3ADSB an amount of enzymes belonging to GH7 family (that includes mainly enzymes from fungi) was detected, although in this sample only a small amount of fungi was identified. The cellobiohydrolases belonging to GH7 family are the most active exoglucanases known (Xu et al., 2007).

The abundance of hemicellulases detected in the three investigated samples was comparable with the percentage occurred in bovine rumen (Brulc et al., 2009) and macropod gut (Pope et al., 2010). In T3ADSB, more that 1% of CAZymes belonged to GH10 family. These enzymes have received much attention for their use in degradation of lignocellulosic biomass for biochemicals production, due to their involvement in breaking down of xylan, the major component of the hemicellulose. Moreover, for the three samples a percentage of 1-2% of enzymes belonging to Glycoside Hydrolases family 28 was identified. These CAZymes are involved in the degradation of pectin, a structural constituent of the plant cell wall.

The profiles of debranching enzymes detected in the three samples showed ~1% of GH 51, higher than detected in yak and cow rumen (Dai et al., 2012; Hess et al., 2011) and snail crop (Cardoso et al., 2012). Moreover, the samples T3ESB and T3PSB revealed a relative abundance of GH 67. The enzymes belonging to these two families are required for the optimal breakdown of glucoronoarabinoxylans (GAXs), one of the major component of hemicellulose, composed by β(1-4)-D-xylose linked polymers branched with arabinose and glucuronic acid. Interestingly, in the samples from *Eucalyptus camaldulensis* and *Populus nigra* 2.5% and 0.6% respectively of total GHs belonging to GH78 α-L-rhamnosidases. These enzymes catalyze the hydrolysis of α-L-rhamnosyl-linkages in L-rhamnosides present in polysaccharides such as rhamnogalacturonan.

Furthermore, the in-depth KEGG pathway mapping of the putative genes coding for Enzyme Commission (EC) number activities related to the polysaccharides

hydrolysis confirmed that all three analyzed samples were a valuable source of a full set of diversified (hemi)cellulases and accessory enzymes required for an effective pretreated lignocellulosic biomass hydrolysis (Hu et al., 2011 & 2015).

Materials and methods

- Lignocellulosic biomasses and environmental DNA extraction

Chipped wood from *A. donax*, *E. camaldulensis* and *P. nigra* was used to form piles of approximately 30 kg that were submitted to biodegradation under natural conditions as previously reported (Ventorino et al., 2015). Briefly, the biomass piles were placed without any coverage under oak trees in the woodland at the Department of Agriculture (Naples, Italy). After 135 days of natural biodagradation, samples of 0.5 kg were collected from the external part (right and left side of the pile) and the internal central part of the biomass, milled and stored at −20 °C until use.

3 g of each milled biomass were used to isolate the total environmental DNA (eDNA), including genetic material from microorganisms adherent to the plant biomass. The eDNA extraction was performed by using the PowerSoil® DNA Isolation Kit (MO BIO Laboratories, INC. CARLSBAD, CA) according to the manufacturer's instructions. NanoDrop and Qubit Fluorometer tests were performed to verify the level of purity of recovered eDNA. About 25 µg of each eDNA samples were sent to BGI Tech Solutions (Hongkong) Co., Limited for further analyses.

- Metagenome shotgun sequencing and assembly

Qualified paired-end libraries were constructed from the three eDNA samples with different clone insert sizes (≤800bp) and then subjected to de novo sequencing by using Illumina HiSeq2000 platform (Illumina, San Diego, CA, USA) following standard pipelines. The generated raw data were pretreated to remove low quality reads: reads with 3 N; reads contaminated by adapter (15 bases overlapped by reads and adapter); reads with 36 low quality bases.

The obtained Clean Data were used to perform the metagenome sequences. Before assembly, k-mer analysis was done to evaluate the sequencing depth for each sample. SOAPdenovo (Version 1.06) (Li et al., 2010) was used to assemble filtered data in contigs and scaffolds and assembly results were optimized by in-house scripts using the SOAP-aligner tool.

- Metagenome analyses

To evaluate the microbial composition the high-quality filtered reads were submitted to SOAPaligner version 2.21 (Li et al., 2009) to obtain a taxonomy assignment using bacteria, fungi and archaea sequences extracted from NCBI NT database. Assembled contigs are used to predict genes by using MetaGeneMark Software (version 2.10, default parameters) based on assembly results.

Functional annotations of predicted amino acid sequences were performed by BLASTP (version 2.2.23). In particular, the metabolism pathway assignment of the predicted protein was performed using the Enzyme Commission (EC) number in the Kyoto Encyclopedia of Genes and Genomes (KEGG) – version 59 – (Kanehisa et al., 2012) databases. and the annotation of each contig with functional categories was carried out by blasting against Evolutionary genealogy of genes: Non-supervised Orthologous Groups（eggNOG) – version 3.0 (Powell et al., 2012). Both comparison were performed by using BLAST (Altschul et al., 1990) with e-value threshold of 1e-5

and a 40% minimum percentage of identity to assign the subject sequence to a specific function family. Moreover, in order to explore in depth the ability of the microbial biodiversity detected in the samples to degrade lignocellulose, the putative encoded protein sequences were first compared to the full length sequences of the CAZy database using BLAST (Altschul et al., 1990) and query sequences that produced a e-value $>10^{-6}$ were discarded. Query sequences that produced a e-value $<10^{-6}$ and aligned over their entire length with a protein in the database with >50% identity were automatically assigned to the same family as the subject sequence. The remaining query sequences were subjected to manual curation which involved BLAST searches against a library built with partial sequences corresponding to individual GH, PL, CE and CBM modules and examination of the conservation of specific family patterns and features such as catalytic residues (where known).

3.2.2 - Function-driven approach of expression library

Results and discussion

In order to further analyze the ability of the above mentioned biomasses subjected to natural biodeterioration to produce novel biocatalysts for lignocellulose conversion, the set up of an expression metagenomic fosmid library was performed. We focused on samples from *Arundo donax* (AD) collected after 135 and 180 days (T3 and T4 respectively) of biodeterioration in filed (C) or underwood (SB) due to the fact that these biodegradation times ensured a high biodiversity with an interesting relative abundance of strains able to degrade lignocellulosic biomasses, as reported in Ventorino et al., 2015. In particular, the cells from samples ADT3C (heaps of *Arundo Donax* put in field at 135 days of natural biodeterioration), ADT3SB (heaps of *Arundo Donax* put in underwood at 135 days of natural biodeterioration), ADT4C (heaps of *Arundo Donax* put in field at 180 days of natural biodeterioration), ADT4SB (heaps of *Arundo Donax* put in underwood at 180 days of natural biodeterioration) were recovered and the metagenomic DNAs were extracted. The sample ADT4C was selected, because of its low contaminations of proteins and solvents (260nm/280nm=1.78; 260nm/230nm=0.82), high concentration (580.9 ng/μL) and size of shared DNA (25-30 kb fragments) in the range of 25-40 kb that are necessary to create a representative library. 1100 was the calculated number of clones required to ensure a 99% probability of a given DNA sequence of *E. coli* (genome = 4.7 Mb) being contained within a fosmid library composed of ~25 kb inserts. Based on the titer of the packaged obtained fosmid clones (~$2x10_8$ cfu/mL), EPI100™ cells were infected with the dilution of the phage particles to obtain a library of 1100 clones. The library was screened for the ability to degrade CMC and xylan in order to detect novel endo-cellulases and xylanases. No clones showed activity halo towards both CMC and xylan after Congo Red assay. Since the high detected biodiversity in the sample showed an abundance of microorganisms producers of enzymes capable of degrading cellulose and hemicellulose, the absence of positive clones in the library may have been caused by a number of factors. As reported in studies most recent than these research activities, a larger number of clones (from 10000 up to 1000000) is required in order to make up for a small number of clones which are active on selective substrates (Kumar et al., 2015) and more effective selective chromogenic media (Bastien et al., 2013), such as AZCL-HE-Cellulose and AZCL-Xylan (Megazyme, Ireland), could be used for the detection of (hemi)cellulases in fosmid metagenomic libraries.

Materials and methods

- Cells recovery and DNA extraction and shearing.

The cells from heaps of *Arundo donax* were recovered according to the methods described by Whitehouse et al. (1994). 500 mg of each recovered cells samples were used to extract the metagenomic DNA by using FastDNA™ SPIN kit for Soil (MP Biomedicals, Solon, Ohio, US) following the supplier's instructions. The DNAs were shared by the Hydroshear® DNA (Digilab, Inc., Marlborough, Massachusetts, USA). The control of the shared DNAs was performed by running the samples on a 20-cm long, 1% agarose gel at 30-35 V overnight and by using NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA).

- Construction of metagenomic fosmid library

Metagenomic 20-30 kb DNA fragments of the selected sample were recovered and used for the fosmid library construction by using the EpiFos™ Library Production Kit (Epicentre Technologies, Madison, Wisconsin, USA), following the manufacturer's indications. The number of fosmid clones required to reasonably ensure that any given DNA sequence is contained within the library was determinated by the formula (Hohn, 1979): $N = \ln(1-P)/\ln(1-f)$
where *P* is the desired probability (expressed as a fraction); *f* is the proportion of the genome contained in a single clone; *N* is the required number of fosmid clones. Based on the titer of the obtained packaged fosmid clones, EPI100™ cells were infected with the proper dilution of the phage particles to obtain the desired numbers of clones. The infected clones were selected on LB plates (primary plates) with 12.5 µg/ml chloramphenicol (fosmid selection marker).

- Functional screening of metagenomic library

For screening purposes, the replica plate technique was performed. The clones from the primary plates were inoculated on secondary agar plates containing as carbon sources CMC (CMC 0,5%; yeast extract 0,7%; Agar 1,5%; KH_2PO_4 0,4%; Na_2HPO_4 0.4%; $MgSO_4$ 0,02%; $FeSO_4$ 0,0004%; $CaCl_2$ 0,0001%) or Xylan (Xylan 0,5%; yeast extract 0,7%; Agar 1,5%; KH_2PO_4 0,4%; Na_2HPO_4 0.4%; $MgSO_4$ 0,02%; $FeSO_4$ 0,0004%; $CaCl_2$ 0,0001%) for the selection of positive clones producing endo-cellulase and xylanase activity respectively. The secondary plates were incubated at 37°C for 72 h. After incubation, the halo of glucan or xylan hydrolysis were revealed by staining the agar replica with 0.1% of Congo red (30 minutes of Congo red incubation at room temperature followed by washing with 5 M NaCl).

Acknowledgemnt

The research activities described in this chapter was supported by a grant from the Ministero dell'Università e della Ricerca Scientifica Industrial research project ''Development of green technologies for production of BIOchemicals and their use in preparation and industrial application of POLImeric materials from agricultural biomasses cultivated in a sustainable way in Campania region - BioPoliS'' PON03PE_00107_1, funded in the frame of Operative National Programme Research and Competitiveness 2007-2013 D. D. Prot. N. 713/Ric. 29/10/201.

Chapter 4

Consolidated BioProcessing

Consolidated BioProcessing (CBP) is a potential system for the production of biofuels and other bioproducts from lignocelluloses, combining (hemi)cellulases production, saccharification of cellulose, and fermentation into a single process, and exploiting the ability of engineered microorganisms to perform these reactions (Amore et al., 2013c). CBP is increasingly recognized as having potential for low-cost production of several target products, since the costs of capital investment, substance and other raw materials, and utilities associated with microbial enzyme production can be avoided (Lynd et al., 2005).The biological approach to CBP employs one genetically engineered host capable of both degrading the lignocellulose and producing a range of different biofuels and other bioproducts, such as fatty acid ethyl esters, pinene, succinic acid and lactate. However, nowadays the main CBP application is related to the bioethanol production (Brethauer and Studer, 2014). Since natural microorganisms with the capability for efficient enzyme production, lignocellulose saccharification, and ethanol production are not currently available, the development of engineered organisms has been pursued by two strategies, native and recombinant cellulolytic strategies. The native cellulolytic strategy (*Category I* CBP) involves the engineering of naturally cellulolytic microorganisms (mainly fungi such as *Trichoderma, Aspergillus, Rhizopus* and *Fusarium* and some thermophilic anaerobes belonging to the genera *Clostridium*) to improve ethanol production. The recombinant cellulolytic strategy (*Category II* CBP) involves the engineering of non-cellulolytic organisms that exhibit high ethanol yield and titer (yeasts such as *Saccharomyces cerevisiae*, *Hansenula polymorpha*, *Kluyveromyces marxianus* and *Scheffersomyces stipitis* and bacteria such as *Escherichia coli*, *Klebsiella oxytoca* and *Zymomonas mobilis*) so that they express a heterologous cellulase system to enable cellulose utilization (Tomohisa and Akihiko, 2012).

4.1 – Development of a *Category II* Consolidated BioProcessing system

The new cellulolytic strain *Streptomyces sp.* G12 was isolated from industrial waste based compost in the laboratory where this PhD project was carried out. Investigating the enzymes responsible for cellulase activity, the endoglucanase CelStrep was identified, recombinantly expressed in *Escherichia coli* and characterized exhibiting a relevant thermoresistance increasing its potential for cellulose conversion (Amore et al., 2012a). Thus, this enzyme was chosen to engineer a *S. cerevisiae* commercial strain in order to develop a one-step consolidated bioprocessing of lignocelluloses to bioethanol by a recombinant cellulolytic strategy (*Category II* CBP). The well-known ethanol producing yeast *S. cerevisiae* is an attractive host for developing a CBP because of its capability of robust fermentative growth and widespread use in industries. However *S. cerevisiae* is unable to hydrolyze polysaccharides. Despite several reports of successful expression of cellobiohydrolases-encoding genes in *S. cerevisiae*, the titres achieved were generally too low to allow CBP.

Results and discussion

The sequence of gene coding for the enzyme CelStrep from *Streptomyces sp.* G12, named *celstrep*, was synthesized with optimized codon usage of *S. cerevisiae* (*cellstrepsacch*) (figure 19) and the leader sequence encoding the signal peptide of *Streptomyces* was replaced by the leader sequence coding for the signal peptide of

Invertase 2 of the yeast (INV2) in order to increase the chance of the secretion of the protein outside of the yeast cells. This replacement was carried out by inserting the cassette INV2-*cellmatsacch* (figure 20).

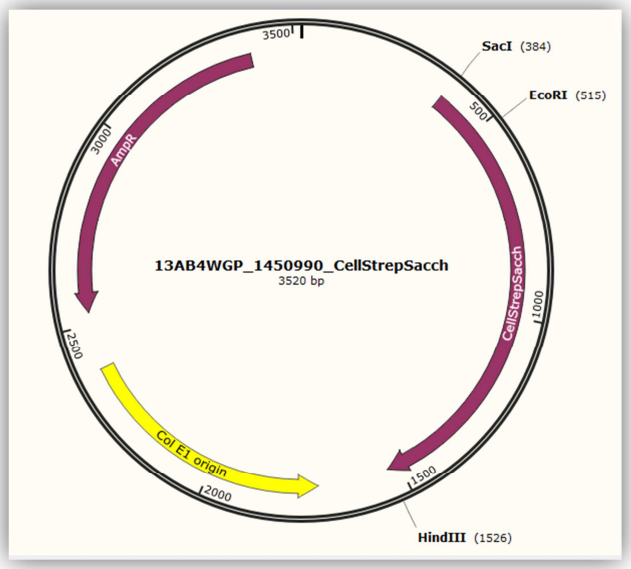


Figure 19 – Map of the synthetic vector containing *cellstrepsacch*.

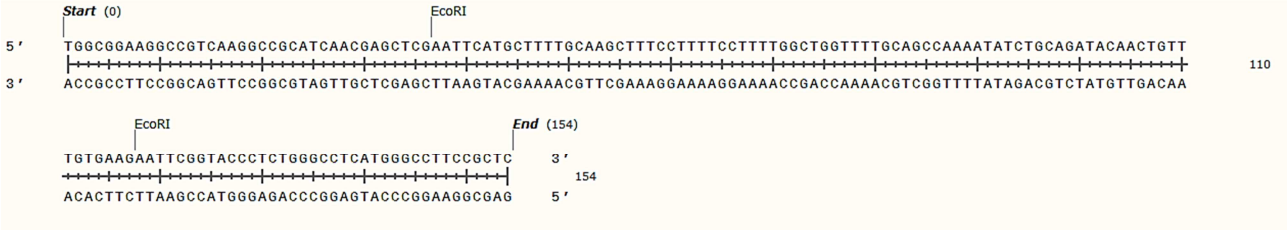


Figure 20 – Sequence of the cassette INV2-*cellmatsacch*.

This tailor made cDNA was cloned in pSal4 plasmid, a shuttle vector containing sequences for the propagation and selection both in bacterial and in yeast cells. The vector pSAL4-INV2-*cellmatsacch* was overexpressed in *E. coli* Top 10 and used to transform *S. cerevisiae* Y294 commercial strain. The transformant *S. cerevisiae* was grown on Synthetic Defined (SD) medium. In order to test the endocellulase activity of the recombinant clones of the yeast, the same colonies were grown on SD plates containing 0.5% CMC: every clone showed an activity halo after Congo Red assay.

Three different clones were grown in three liquid media with different carbon sources: SD, CMC 0,5% and YP-CMC 2% (by testing the use of CMC as inducer for the endo-cellulase production), supplemented or not with 0.6 mM copper sulfate. For each medium, the time-course of cells growth was daily evaluated and the endo.cellulase activity was detected both on extracellular broth and on intracellular material, after mechanical cell lysis. The obtained data showed that an acceptable growth rate (0.02 h^{-1}) was achieved only in SD medium and no clones showed detectable extra- and intra-cellulase activity. These results indicated that this engineered strain cannot grow on cellulose as self-catalyzing cellulolytic microorganisms, possibly because of poor active cellulase expression. This is one of the major drawback of the *Category II* CBP. In fact, as reported by Zhang and Zhang, 2013, to achieve the self-supporting growth based on recombinant cellulases, the

required cellulase expression levels is 1-10% of cellular protein. To overcome this bottleneck, several studies were carried out to co-express different cellulases in a unique host in order to enhance the cellulases expression levels (Fujita et al., 2004; Den Haan et al., 2007). The obtained results showed that the engineered *S. cerevisiae* Y294 with the unique cellulase CelStrep does not allow to obtain the minimum required endo-cellulase expression levels and therefore it is a not adequate system for the development of an effective *Category II* CBP.

Materials and methods

- Strains, media and plasmids

The *Escherichia coli* strain Top 10 (F-mcrA D (mrrhsdRMS-mcrBC) f80lacZDM15 DlacX74 deoR recA1 araD139 D (ara-leu) 7697 galU galK rpsL (StrR) endA1 nupG) was used in all DNA manipulations. *E. coli* was grown in Luria–Bertani (LB) medium (in grams per liter: 10 bacto tryptone, 10 NaCl, and 5 yeast extract) supplemented, when required, with 100 μg ml^{-1} of ampicillin.

The *Saccharomyces cerevisiae* strain used for heterologous expression was the commercial strain Y294 (MATalpha leu2-3 leu2-112 ura3-52 his3deltatrp1 GAL+ [cir+] - ATCC® Catalog No. 201160™). The *S. cerevisiae* Y294 host strains were cultivated in YEPD medium (10 g l^{-1} yeast extract, 20 g l^{-1} peptone and 20 g l^{-1} glucose) on a rotary shaker at 200 rpm at 28°C. Yeast transformants were selected and maintained on Synthetic Defined (SD) medium (1,5% agar; 0.67% nitrogen base w/o amino acids; 0.5% casaminoacids; 2% galactose; 0.003% adenine; 0.004% tryptophane; 0.008% succinate; 600μM $CuSO_4$) without uracile. Aerobic cultivation was performed in 125 ml Erlenmeyer flasks containing 20 ml SD medium on a rotary shaker at 200 rpm at 28°C.

The shuttle-expression vector pSAL4 was used both for overexpression in *Escherichia coli* Top 10 and for heterologous expression in *Saccharomyces cerevisiae* Y294. pSAL4 is similar to pSAL1 (Mascorro-Gallardo et al., 1996), but it has the pRS426 backbone (Christianson et al., 1992) with the 2 μ origin of replication and the *URA3* marker.

- cDNA synthesis

The gene *celstrepsacch* with optimized codon usage of *S. cerevisiae* and the cassette INV2-*cellmatsacch* were synthesized by Invitrogen™ Corporation (Carlsbad, California, US). The cassette INV2-*cellmatsacch* consists of the leader sequence encoding the signal peptide of *Invertase 2* of yeast and the *cellstrepsacch* sequence up to EcoRI restriction site without the leader sequence encoding the original signal peptide.

- DNA manipulation and vectors construction

Standard protocols were followed for DNA manipulation with enzymes for restriction digests and ligations sourced from Promega Corporation (Fitchburg, Wisconsin, United States). Where applicable, DNA was eluted from agarose gels. The *cellstrepsacch* gene was subcloned as an SacI-HindIII fragment into the corresponding sites of plasmid pSAL4, under transcriptional control of the copper-inducible CUP1 promoter, yielding plasmid pSAL4-*cellstrepsacch*. The cassette INV2-*cellmatsacch* was subcloned as an EcoRI-EcoRI fragment into the corresponding sites of plasmid pSAL4-*cellstrepsacch,* yielding plasmid pSAL4-INV2-

cellmatsacch. The presence of the correct sequence in the plasmid pSAL4-INV2-*cellmatsacch* was verified by sequencing by PRIMM srl (Italy).

- Yeasts transformation and selection

The host strain *S. cerevisiae* Y294 was transformed with the recombinant plasmids by using lithium acetate method with subsequent selection on SD plates. At least three independent transformants were chosen to test their endocellulase activity.

- Endocellulase assays

For qualitative assays, recombinant *S. cerevisiae* strains were cultured on SD plates containing 0.5% CMC for 72 hours at 28°C. After incubation, zones of glucan hydrolysis were revealed by staining the agar replica with 0.1% of Congo red (30 minutes of Congo red incubation followed by washing with 5 M NaCl).

For quantitative assays, yeast transformants were cultured in 125 ml Erlenmeyer flasks up to 10 days with agitation at 200 rpm at 28°C in 20 ml of three different liquid media: SD, CMC 0,5% (CMC 0,5%; yeast extract 0,7%; KH2PO4 0,4%; Na2HPO4 0.4%; MgSO4 0,02%; FeSO4 0,0004%; CaCl2 0,0001%) and YP-CMC 2% (1% Bacto Yeast Extract; 2% Bacto Difto Peptone; 2% CMC) supplemented or not with 0.6 mM copper sulfate. The supernatant was harvested and the extracellular *endo*-1,4-ß-glucanase activity was determined. Moreover, the cells were lysed by using French pressure cell press in order to detect the intracellular *endo*-1,4-ß-glucanase activity. Both extra- and intra-cellular *endo*-1,4-ß-glucanase activity were evaluated by AZO-CMCase assay (Megazyme, Ireland), following the supplier's instructions. The analytical determinations correspond to the mean value of three replicates.

Acknowledgemnt

The research activities described in this chapter was supported by a grant from the Ministero dell'Università e della Ricerca Scientifica Industrial Research Project "Integrated agro-industrial chains with high energy efficiency for the development of eco-compatible processes of energy and biochemical production from renewable sources and for the land valorization - EnerbioChem" PON01_01966, funded in the frame of Operative National Programme Research and Competitiveness 2007–2013 D. D. Prot. n. 01/Ric. 18.1.2010.

Concluding remarks

This phD project was aimed at the valorization of selected lignocellulosic biomasses as source of both fermentable sugars and novel biocatalysts for the production of biobased products via fermentation.

Two different lignocellulosic biomasses (the perennial crop *Arundo donax* and the Newspaper Waste – NW – fraction of Municipal Solid Waste) were tested as source of monosaccharides by enzymatic hydrolysis after Ammonia Fiber Expansion (AFEX) or Extractive Ammonia (EA) pretreatment. The ability of the recombinant endocellulase rCelStrep, α-L-arabinofuranosidase rPoAbf and its evolved variant rPoAbf F435Y/Y446F to improve the saccharification yields was evaluated. In particular, a mixture of purified (hemi)cellulases was chosen as reference and the three enzymes were replaced to the corresponding enzymatic activities in the mix. For the AFEX pretreated *A. donax*, the use of rPoAbf F435Y/Y446F led to obtain a glucan, xylan and arabinan conversion after 72 h of around 62, 63 and 80 % respectively, similar or higher than those (44, 66 and 55 %) achieved by 72 h hydrolysis with the commercial enzymes Novozymes Cellic®, Ctec3 and Htec3. The enzymes rPoAbf, rPoAbf F435Y/Y446F and rCelStrep were also investigated for their effect on the hydrolysis yields by their addition to the commercial enzyme mixture Novozymes Cellic®, Ctec3 and Htec3. The addition of rPoAbf F435Y/Y446F enhanced both xylan and arabinan conversions during the AFEX-pretreated *A. donax* saccharification, achieving 80% after 6 days of hydrolysis. The total polysaccharides conversion yield reached 37.32% for AFEX pretreated NW by adding rPoAbf to the mix whilst the maximum sugars conversion yield for EA pretreated NW was achieved 40.80 % by adding rCelStrep. The maximum glucan conversion yield for EA pretreated NW (45.61%) was obtained by adding rCelStrep to the commercial mix. This value was higher than or comparable to those reported in recent manuscripts adopting hydrolysis conditions similar to those used in this study.

The microbial diversity of natural ecosystems of three biomasses (chipped wood of *Arundo donax, Eucalyptus camaldulensis* and *Populus nigra*) subjected to natural biodegradation in underwood or open field was evaluated by both culture-dependent and culture-independent approaches in order to identify novel lignocellulose-degrading enzymes. One *Pediococcus acidilactici* strain and five *Streptomyces* strains were identified as producers of novel endo-xylanase(s) and -cellulases, respectively, by using traditional microorganisms cultivation based methods. The identification of oligopeptides from the protein(s) involved in the hydrolysis of xylan produced by the *P. acidilactici* strain was performed by a zymographic approach combined with proteomic analysis. The fact that none of these peptides matched with proteins from *P. acidilactici* encourage further researches on this bacterial strain with the aim of the identification and characterization of novel biocatalysts involved in the hemicellulose degradation.
The cellulases from the *Streptomyces* strains were tested in the hydrolysis of pretreated *A. donax*, by the substitution of the corresponding enzymatic activity in a commercial mix chosen as reference. Interestingly, the glucose and xylose yields obtained by the use of the cellulase(s) from the AE-T0-58P (10) strain (4.47±0.5 g/L and 5.87±0.2 g/L respectively after 72 h) were 82% and 85% respectively of the corresponding values obtained by using the reference commercial mixture. This comparison showed that the endo-cellulase activity produced by this strain was a good candidate to replace commercial cellulose mix for the *A. donax* saccharification with satisfactory conversion yields.

Moreover, the genetic material from microorganisms adherent to the abovementioned biomasses – after 135 days of biodeterioration in underwood – were submitted to metagenomic approach. The functional clustering of the open reading frames (ORFs) predicted by the sequencing of the extracted DNAs showed a prevalence of poorly characterized genes belonging to S (function unknown) or R (general function prediction only) eggNOG (evolutionary genealogy of genes Non-supervised Orthologous Groups) category, suggesting the three detected biomasses as potential sources of not yet known genes. 1792, 1279 and 2113 putative Carbohydrate-Active Enzymes (CAZymes) were identified in the samples from *A. donax, E. camaldulensis* and *P. nigra* respectively, corresponding to 1.15%, 0.59% and 3.44% of the total ORFs. However, most of the detected CAZymes in the three samples were involved in hydrolysis and/or rearrangement of glycosidic bonds. In particular, a number of 1059 in *A. donax* (corresponding to 59.10% on total CAZymes and to 0.68% on total ORFs detected), 750 in *E. camaldulensis* (corresponding to 58.60% on total CAZymes and to 0.34% on total ORFs detected) and 1136 in *P. nigra* (corresponding to 53.80% on total CAZymes and to 1.85% on total ORFs detected) predicted proteins were classified as Glycoside Hydrolases (GHs). Interestingly, the GHs abundance in the sample from *P. nigra* (1.85% on total ORFs) was higher than that detected in microbiota of different organisms – such as plant-fed elephant, invasive snail, cow, termite – generally showing several predicted genes involved in hydrolysis of the glycosidic linkage. An in-depth KEGG pathway mapping was carried out for the putative genes coding for enzymes involved in the hydrolysis of glycosidic bonds in complex sugars. The results showed a high percentage of several cellulases (mainly β-glucosidases and endo-1,4-β-glucanases), different hemicellulases and accessory enzymes (mannanases, polygalacturonases and feruloyl esterases), confirming that the three analyzed samples were a reservoire of a full set of diversified biocatalysts required for an effective lignocellulose saccharification.

References

Abbott, D. W., Eirín-López, J. M. & Boraston, A. B. Insight into ligand diversity and novel biological roles for family 32 carbohydrate-binding modules. *Mol. Biol. Evol.* **25**, 155-167 (2008).
Adsul, M. G., Singhvi, M. S., Gaikaiwari, S. A. & Gokhale D. V. Development of biocatalysts for production of commodity chemicals from lignocellulosic biomass. *Bioresour. Technol.* **102,** 4340-4312 (2011).
Allgaier, *et al.*, Targeted Discovery of Glycoside Hydrolases from a Switchgrass-Adapted Compost Community. *PLoS ONE* **5**(1): e8812 (2010).
Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215,** 403-410 (1990).
Anand, A. A. *et al*. Isolation and characterization of bacteria from the gut of Bombyx mori that degrade cellulose, xylan, pectin and starch and their impact on digestion. J*. Insect Sci. 10,* 107 (2010).
Amore, A. *et al.* Cloning and recombinant expression of a cellulase from the cellulolytic strain Streptomyces sp. G12 isolated from compost. *Microb. Cell Fact.* **11,** 164 (2012a).
Amore, A., Pepe, O., Ventorino, V., Birolo, L., Giangrande, C. & Faraco, V. Cloning and recombinant expression of a cellulase from the cellulolytic strain Streptomyces sp. G12 isolated from compost. *Microb Cell Fact.* **11** (2012).
Amore, A., Pepe, O., Ventorino, V., Aliberti, A. & Faraco, V. Cellulolytic Bacillus strains from natural habitats - A review. *Chim. Oggi/Chem. Today* **31,** 49–52 (2013a).
Amore, A. *et al.* Industrial waste based compost as a source of novel cellulolytic strains and enzymes. *FEMS Microbiol. Lett.* **339,** 93-101 (2013b).
Amore, A., Giacobbe, S. & Faraco, V. Consolidated Bioprocessing for Improving Cellulosic Ethanol Production. In: Lignocellulose Conversion. V. Faraco (ed). Springer. (2013c)
Arfi, Y. *et al*. Characterization of salt-adapted secreted lignocellulolytic enzymes from the mangrove fungus *Pestalotiopsis* sp. *Nat. Commun*. **4,** 1810 (2013).
Bailey, M.J., Biely, P. & Poutanen, K. Interlaboratory testing of methods for assay of xylanase activity. *J Biotechnol.* **3,** 257-270 (1992).
Balan, V., Bals, B., Chundawat, S.P., Marshall, D. & Dale B.E. Lignocellulosic biomass pretreatment using AFEX. *Methods Mol Biol.* **581,** 61-77 (2009).
Bastien, G. *et al.* Mining for hemicellulases in the fungus-growing termite *Pseudacanthotermes militaris* using functional metagenomics. *Biotechnol Biofuels.* **6** (2013).
Beguin, P. Detection of cellulase activity in polyacrylamide gels using Congo red-stained agar replicas. *Anal Biochem.* **131,** 333–336 (1983).
Beeson, W. T., Vu, V. V., Span, E. A., Phillips, C. M. & Marletta, M. A. Cellulose degradation by Polysaccharide Monooxygenases. *Ann. Rev. Biochem.* **84,** 923–46 (2015).
Binod, P., Janu. K.U., Sindhu, R. & Pandey, A. Hydrolysis of Lignocellulosic Biomassfor Bioethanol Production. In: Biofuels – alternative feedstock and conversion processes. Editors: Biofuels, Burlington: Academic Press, 229-250 (2011).
Brethauer, S. & Studer, M.H. Consolidated bioprocessing of lignocellulose by a microbial consortium. *Energy Environ Sci.* **7,** 1446-1453 (2014).
Brulc, J. M. *et al.*, Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. *Proc. Natl. Acad. Sci. USA* **106** (6), 1948-1953 (2009).

Cahyani, V. R., Matsuya, K., Asakawa, S. & Kimura, M. Succession and phylogenetic profile of eukaryotic communities in the composting process of rice straw estimated by PCR-DGGE analysis. *Biol. Fertil. Soils* **40,** 334–344 (2004).
Camassola, M. & Dillon, A. J. P. Biological pretreatment of sugar cane bagasse for the production of cellulases and xylanases by *Penicillium echinulatum*. *Indus. Crops Prod.* **29**, 642–647 (2008).
Cantarel, B. L*. et al.* The Carbohydrate-Active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Res.* **37,** D233-D238 (2009).
Cardoso, A. M. *et al.*, Metagenomic Analysis of the Microbiota from the Crop of an Invasive Snail Reveals a Rich Reservoir of Novel Genes. *PLoS ONE* **7** (11): e48505 (2012).
Castillo, J. M., Romero, E. & Nogales, R. Dynamics of microbial communities related to biochemical parameters during vermicomposting and maturation of agroindustrial lignocellulose wastes. *Bioresour. Technol.* **146,** 345–354 (2013).
Chen H. Y., Xue D. S., Feng X. Y. & Yao S. J. Screening and production of ligninolytic enzyme by a marine-derived fungal *Pestalotiopsis* sp. J63. *Appl Biochem Biotechnol* **165,** 1754-1769 (2011).
Christianson, T.W., Sikorski, R.S., Dante, M., Shero, J.H. & Hieter, P. Multifunctional yeast high-copy number shuttle vectors. *Gene.* **110,** 119–122 (1992).
Chu, K.H. & Feng, X. Enzymatic conversion of newspaper and office paper to fermentable sugars. *Proc Safety Environm Protection,* **91,** 123–130 (2013).
Contreras Sánchez-Matamoros, R., Serrano, A.M.G., Tejero-Mateo, P., Ollero, J., Saavedra E.M. & Rodríguez-Carvajal, M.A. Structure of the O-antigen of the lipopolysaccharide isolated from Pantoea ananatis AEP17, a rhizobacterium associated with rice. *Carb Res,* **369,** 25-30 (2013).
Crawford, D.L. Biodegradation of agricultural and urban wastes. In: Actinomycetes in biotechnology, M.Goodfelow (Ed.), Academic Press, London. 433–459 (1988).
da Costa Sousa *et al.* Next-Generation Ammonia Pretreatment Enhances Biofuel Production from Biomass via Simultaneous Cellulose Decrystallization and Lignin Extraction. *Energy and Environmental Science.* Accepted. (2016)
Dai, X. *et al.*, Metagenomic Insights into the Fibrolytic Microbiome in Yak Rumen. *PLoS ONE* **7** (7): e40430 (2012).
DeAngelis, K. M. *et al.* Strategies for enhancing the effectiveness of metagenomic-based enzyme discovery in lignocellulolytic microbial communities. *Bioenerg. Res.* **3,** 146–158 (2010).
Den Haan, R., Rose, S.H., Lynd, L.R. & Van Zyl W.H. Hydrolysis and fermentation of amorphous cellulose by recombinant *Saccharomyces cerevisiae Metab. Eng.* **9,** 87–94 (2007).
Dimarogona, M., Topakas, E.& Christakopoulos, P. Recalcitrant polysaccharide degradation by novel oxidative biocatalysts. *Appl Microbiol Biotechnol.* **97,** 8455–65 (2013).
Diodato, N., Fagnano, M. and Alberico, I. CliFEM–climate forcing and erosion response modelling at long-term Sele River research basin (southern Italy). *Nat Hazards Earth Syst Sci.* **9,** 1693–1702 (2009).
Diodato, N., Fagnano, M. and Alberico, I. Geospatial–and–visual modeling for exploring sediment source areas across the Sele river landscape, Italy. *Ital J Agron.* **6,** 85–92 (2011).
Di Pasqua, R. *et al*. Influence of different lignocellulose sources on endo-1,4-ß-glucanase gene expression and enzymatic activity of *Bacillus amyloliquefaciens* B31C. *Bioresources* **9,** 1303-1310 (2014).

Do, T. H. *et al.*, Mining biomass-degrading genes through Illumina-based de novo sequencing and metagenomic analysis of free-living bacteria in the gut of the lower termite Coptotermes gestroi harvested in Vietnam. *J. Biosci. Bioeng.* **118** (6) 665-671 (2014).
Dougherty, M. J. *et al.* Glycoside hydrolases from a targeted compost metagenome, activity-screening and functional characterization. *BMC Biotechnol*. **12,** 38 (2012).
Ducros, V. *et al.* Substrate Specificity in Glycoside Hydrolase Family 10. *J Biol Chem* **28,** 23020-23026 (2000).
Eggeman, T. & Elander, R.T. Process and economic analysis of pretreatment technologies. *Biores Technol.* **96,** 2019–2025 (2005).
Eida, M. F., Nagaoka, T., Wasaki, J. & Kouno, K. Isolation and characterization of cellulose-decomposing bacteria inhabiting sawdust and coffee residue composts. *Microbes Environ.* **27,** 226–233 (2012).
Epicoco, M., Oltra, V. and Saint Jean, M. Knowledge dynamics and sources of eco-innovation: Mapping the Green Chemistry community. *Technol Forecast Soc.* **81,** 388-402 (2014).
Fagnano, M., Diodato, N., Alberico, I. and Fiorentino, N. An overview of soil erosion modelling compatiblewith RUSLE approach. *Rend. Fis. Acc. Lincei.* **23**, 69-80 (2012).
Fagnano, M., Impagliazzo, A., Mori, M. and Fiorentino N. Agronomic and environmental impacts of giant reed (Arundo donax L.): results from a long-term field experiment in hilly areas subject to soil erosion. *Bioenergy Res.* **8,** 415–422 (2015).
Feng, Y. *et al*. Degradation of raw corn stover powder (RCSP) by an enriched microbial consortium and its community structure. *Bioresour. Technol.* **102,** 742-747 (2011).
Fiorentino, N. *et al.* Biomass accumulation and heavy metal uptake of giant reed on polluted soil in southern Italy. *J Biotechnol.* **150,** 261 (2010).
Fiorentino, N. *et al.* Assisted phytoextraction of heavy metals: compost and *Trichoderma* effects on giant reed (*Arundo donax* L) uptake and soil N-cycle microflora. *Ital. J. Agron.* **8,** 244-254 (2013).
Forte, A., Zucaro, A., Fagnano, M., Bastianoni, S., Basosi, R. & Fierro A. LCA of Arundo donax L. lignocellulosic feedstock production under Mediterranean conditions: a three-year study. *Biomass Bioenergy*. **73,** 32–47 (2015).
Fu, X. *et al.* A novel endoglucanase (Cel9P) from a marine bacterium Paenibacillus sp. BME-14. *Appl. Biochem. Biotechnol.* **160,** 1627-1636 (2010).
Fujita, Y., Ito, J., Ueda, M., Fukuda, H. & Kondo, A. Synergistic saccharification, and direct fermentation to ethanol, of amorphous cellulose by use of an engineered yeast strain codisplaying three types of cellulolytic enzyme. *Appl Environ Microbiol,* **70,** 1207–1212 (2004).
Gao, D., Chundawat, S. P. S., Krishnan, C., Balan, V. & Dale, B. E. Mixture optimization of six core glycosyl hydrolases for maximizing saccharification of ammonia fiber expansion (AFEX) pretreated corn stover. *Bioresour. Technol*. **101,** 2770–2781 (2010).
Gao, D. *et al.* Hemicellulases and auxiliary enzymes for improved conversion of lignocellulosic biomass to monosaccharides. *Biotechnol. Biofuels*. **4,** 5 (2011).
García-Fraile, P. *et al. Rhizobium cellulosilyticum* sp. nov., isolated from sawdust of *Populus alba*. *Int. J. Syst. Evol. Microbiol*. **57,** 844-848 (2007).
Gagne-Bourgue, F., Aliferis, K., Seguin, P., Rani, M., Samson, R. & Jabaji, S. Isolation and characterization of indigenous endophytic bacteria associated with leaves of switchgrass (Panicum virgatum L.) cultivars. *J Appl Microbiol.* **114,** 836-853 (2012).

Genome Online Database, Available at https://gold.jgi.doe.gov/.
Gherbawy, Y. A. M. H. & Abdelzaher, H. M. A. Isolation of fungi from tomato rhizosphere and evaluation of the effect of some fungicides and biological agents on the production of cellulase enzymes by *Nectria haematococca* and *Pythium ultimum* var. *ultimum*. *Czech Micol.* (1999).
Giacobbe, S., Vincent, F. & Faraco, V. Development of an improved variant of GH51 α-l-arabinofuranosidase from Pleurotus ostreatus by directed evolution. *N Biotechnol.* **31,** 230-236 (2014).
Giacobbe, S., Balan, V., Montella, S., Fagnano, M., Mori, M. & Faraco, V. Assessment of bacterial and fungal (hemi)cellulose-degrading enzymes in saccharification of ammonia fibre expansion-pretreated Arundo donax. *Appl Microbiol Biotechnol.* (2015)
Goedegebuur, F. *et al.* Cloning and relational analysis of 15 novel fungal endoglucanases from family 12 glycosyl hydrolase. *Curr. Genet.* **41,** 89-98 (2002).
Haberl, H., Beringer, T., Bhattacharya, S. C., Erb, K-H & Hoogwijk, M. The global technical potential of bio-energy in 2050 considering sustainability constraints. *Curr. Opin. Environ. Sustain.* **2,** 394-403 (2011).
Hamid, R. *et al.* Chitinases: An update. *J. Pharm. Bioallied Sci.* **5**, 21-29 (2013).
Hermann, B. G., Blok, K. and Patel, M. K. Producing bio-based bulk chemicals using industrial biotechnology saves energy and combats climate change. *Environ Sci Technol.* **41**, 7915-21 (2007).
Hess, M. *et al.* Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. *Science* **331,** 463-467 (2011).
Heux, S., I. Meynial-Salles, I., O'Donohue M.J. and Dumon C. White biotechnology: State of the art strategies for the development of biocatalysts for biorefining. *Biotechnol Adv.***33,** 1653–1670 (2015).
Hofvendahl, K. & Hahn–Hagerdal, B. Factors affecting the fermentative lactic acid production from renewable resources. *Enz Microb Technol.* **26,** 87-107 (2000).
Hohn, B. In vitro packaging of lambda and cosmid DNA. *Methods Enzymol.* **68,** 299-309 (1979).
Hu, J., Arantes, V. & Saddler, J. N., The enhancement of enzymatic hydrolysis of lignocellulosic substrates by the addition of accessory enzymes such as xylanase: is it an additive or synergistic effect? *Biotechnol. Biofuels* **4** (36) (2011).
Hu, J. *et al.* The addition of accessory enzymes enhances the hydrolytic performance of cellulase enzymes at high solid loadings. *Bioresour. Technol.* **186**, 149–153 (2015).
Ilmberger, N. *et al.* A comparative metagenome survey of the fecal microbiota of a breast- and a plant-fed Asian elephant reveals an unexpectedly high diversity of glycoside hydrolase family enzymes. *PLoS ONE* **9** (9): e106707 (2014).
Ioelovich, M. Recent findings and the energetic potential of plant biomass as a renewable source of biofuels – a review. *Biores.* **10** (2015).
Jensen, L. J. *et al.* eggNOG: automated construction and annotation of orthologous groups of genes. *Nucleic Acids Res.* **36, D250-254** (2008).
Jones, S. M., Van Dyk, J. S. & Pletschke, B. I. *Bacillus subtilis* SJ01 produces hemicellulose degrading multi-enzyme complexes. *Bioresources* **7,** 1294–1309 (2012).
Jorgensen, H., Eriksson, T., Borjesson, J., Tjerneld, F. & Olsson, L. Purification and characterization of five cellulases and one xylanase from *Penicillium brasilianum* IBT 20888. *Enzyme Microb. Technol.* **32,** 851–61 (2003).

Kandasamy S. *et al.* Screening and Identification Of Antibiotic Producing Actinomycetes And Their Antagonistic Activity Against Common Pathogens. *World Res J Antimicrob Agents.* **1,** 7-10 (2012).
Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. & Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* **40,** D109–D114 (2012).
Kanokratana, P. *et al.* Insights into the phylogeny and metabolic potential of a primary tropical peat swamp forest microbial community by metagenomic analysis. *Microb. Ecol.* **61,** 518-528. (2011).
Kim, S.B., Kim, H.J. & Kim, C.J. Enhancement of the enzymatic digestibility of waste newspaper using Tween. *Appl Biochem Biotechnol*. **129-132,** 486-495 (2006).
Kim, J.S., Lee, Y.Y. & Kim, T.H. A review on alkaline pretreatment technology for bioconversion of lignocellulosic biomass. *Bioresour Technol.* **199,** 42-8 (2016).
Krogh, K. B. R. M. *et al.* Characterization and kinetic analysis of a thermostable GH3 b-glucosidase from *Penicillium brasilianum*. *Appl. Microbiol. Biotechnol*. **86,** 143–154 (2010).
Kuhad, R.C., Mehta, G., Gupta, R. & Sharma, K.K. Fed batch enzymatic saccharification of newspaper cellulosics improves the sugar content in the hydrolysates and eventually the ethanol fermentation by *Saccharomyces cerevisiae*. *Biomass and Bioene*rgy **34,** 1189–1194 (2010).
Kumar, S., Krishnani, K, Bhushan, B. & Brahmane, M.P. Metagenomics: Retrospect and Prospects in High Throughput Age. *Biotechnology Research International.* (2015)
Lednicka, D., Mergaer, J., Cnockaer M.C. & Swings, J. Isolation and Identification of Cellulolytic Bacteria Involved in the Degradation of Natural Cellulosic Fibres. *System Appl Microbiol.* **23,** 292-299 (2000).
Laemmli, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* **227,** 680–685 (1970).
Li, L. L., McCorkle, S. R., Monchy, S., Taghavi, S. & van der Lelie, D. Bioporspecting metagenomes: glycosyl hydrolases for converting biomass. *Biotechnol. Biofuels* **18** (2) (2009).
Li, R. *et al.* SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* **25,** 1966-1967 (2009).
Li, R. *et al.* De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* **20,** 265-272 (2010).
Li, H. *et al.* Molecular analyses of the functional microbial community in composting by PCR-DGGE targeting the genes of the b-glucosidase. *Bioresour. Technol.* **134,** 51-58 (2013).
Lichtenthaler, F.W. and Peters, S. Carbohydrates as green raw materials for the chemical industry. *C. R. Chimie.* **7**, 65-90 (2004).
Liguori, R., Amore, A. & Faraco, V. Waste valorization by biotechnological conversion into added value products. *Appl. Microbiol. Biotechnol*. **97,** 6129-6147 (2013).
Liguori, R., Soccol C. R., Porto de Souza Vandenberghe, L., Woiciechowski, A.L., Faraco, V. Second Generation Ethanol Production from Brewers' Spent Grain. *Energies.* **8,** 2575-2586 (2015).
Liguori, R., Ventorino, V., Pepe, O. & Faraco, V. Bioreactors for lignocellulose conversion into fermentable sugars for production of high added value products. *Appl. Microbiol. Biotechnol*. **100,** 597-611 (2016).

Lopez-Gonzalez, J. A. *et al*. Dynamics of bacterial microbiota during lignocellulosic waste composting: studies upon its structure, functionality and biodiversity. *Bioresour. Technol*. **175,** 406-416 (2015).
Lynd, L., van Zyl, W., McBride, J. & Laser, M. Consolidated bioprocessing of cellulosic biomass: an update. *Curr Opin Biotechnol,* **16,** 577-83 (2005).
Mariani, C. *et al*. Origin, diffusion and reproduction of the giant reed (*Arundo donax* L.): a promising weedy energy crop. *Ann. Appl. Biol*. **157,** 191-202 (2010).
Mascorro-Gallardo, J.O., Covarrubias, A.A. & Gaxiola, R. Construction of a *CUP1* promoter-based vector to modulate gene expression in *Saccharomyces cerevisiae*. *Gene.* **172,** 169–170 (1996).
Montella, S., Amore, A. & Faraco, V. Metagenomics for the development of new biocatalysts to advance lignocellulose saccharification for bioeconomic development. *Crit. Rev. Biotechnol*. **18,** 1-12 (2015).
Morales-Jiménez, J., Zúñiga, G., Ramírez-Saad, H.C. & Hernández-Rodríguez, C. Gut-associated bacteria throughout the life cycle of the bark beetle Dendroctonus rhizophagus Thomas and Bright (Curculionidae: Scolytinae) and their cellulolytic activities. *Microb. Ecol.* **64,** 268-278 (2012).
Mussatto, S. I. & Teixeira, J. A. Lignocellulose as raw material in fermentation processes. In: Méndez-Vilas A (ed) Current research, technology and education topics in applied microbiology and microbial biotechnology. **2,** 897–907 (2010).
Narayanaswamy, N., Dheeran, P., Verma, S. & Kumar, S.. Biological pretreatment of lignocellulosic biomass for enzymatic saccharification. In: Pretreatment Techniques for Biofuels and Biorefineries, ed Fang Z., editor. (Berlin; Heidelberg: Springer Verlag;), 3–34 (2013).
nee' Nigam, P. S., Nutan Gupta, N. & Anthwal, A. Pre-treatment of Agro-Industrial Residues. In: Biotechnology for Agro-Industrial Residues Utilisation. Springer. 13-33 (2009).
Olofsson, K., Bertilsson, M. & Lidén, G. A short review on SSF - an interesting process option for ethanol production from lignocellulosic feedstocks. *Biotechnol Biofuels*. **1** (2008).
Pastor, F. J. *et al*. Molecular cloning and characterization of a multidomain endoglucanase from *Paenibacillus* sp BP-23: evaluation of its performance in pulp refining. *Appl. Microbiol. Biotechnol.* **55,** 61-68 (2001).
Pope, P. B. *et al.*, Adaptation to herbivory by the Tammar wallaby includes bacterial and glycoside hydrolase profiles different from other herbivores. *PNAS* **107**, 14793-14798 (2010).
Powell, S. *et al*. eggNOG v3.0: orthologous groups covering 1133 organisms at 41 different taxonomic ranges. *Nucleic Acids Res.* **40,** D284–D289 (2012).
Powell, *et al.*, eggNOG v4.0: nested orthology inference across 3686 organisms. *Nucleic Acids Res.* **42** (2013).
Robledo, M*. et al. Rhizobium* cellulase CelC2 is essential for primary symbiotic infection of legume host roots. *Proc. Natl. Acad. Sci. USA* **105,** 7064-7069 (2008).
Ryckeboer, J., Mergaert, J., Coosemans, J., Deprins, K. & Swings, J. Microbiological aspects of biowaste during composting in a monitored compost bin. *J. Appl. Microbiol*. **94,** 127-137 (2003).
Sambasiva, R.K.R, Tripathy, N.K., Mahalaxmi, Y. & Prakasham R.S. Laccase- and peroxidasefree tyrosinase production by isolated microbial strain. *J Microbiol Biotechnol.*; **22,** 207–214 (2012).
Scheidel, A. and Sorman, A.H. Energy transitions and the global land rush: Ultimate drivers and persistent consequences. *Glob Environ Change*. **22,** 588–595 (2012).

H. Pooja Sharma, K. Mukesh, V. Mensi, K. Kirti, M. Shweta and K. R. Jitendra, "From bacterial
genomics to metagenomics: concept, tools and recent advances.," Indian J Microbiol, no. 48, pp.
173-194, 2008.
Sharma, P., Mukesh, K., Mensi, V., Kirti, K., Shweta M. & Jitendra, K.R. From bacterial genomics to metagenomics: concept, tools and recent advances. *Indian J Microbiol.* **48,** 173-194 (2008).
Solecka, J., Zajko, J., Postek, M. & Rajnisz, A. Biologically active secondary metabolites from Actinomycetes. *Cent Eur J Biol.* **7,** 373-390 (2012).
Stephen, J.D., Mabee, W.E. & Saddler, J.N. Will second-generation ethanol be able to compete with first-generation ethanol? Opportunities for cost reduction, *Biofuels, Bioprod. Biorefin.* **6,** 159-176 (2012).
Subhedar, P.B., Babu, N.R. & Gogate, P.R. Intensification of enzymatic hydrolysis of waste newspaper using ultrasound for fermentable sugar production. *Ultrason Sonochem.* **22,** 326-32 (2015).
Talia, P. *et al*. Biodiversity characterization of cellulolytic bacteria present on native chaco soil by comparison of ribosomal RNA genes. *Res. Microbiol*. **163,** 221-232 (2012).
Taherzadeh, M.J. & Karimi, K. Acid based hydrolysis process for bioethanol production from lignocellulosic materials: a review. *BioRes*. **2,** 472–499 (2007).
Tomohisa H. & Akihiko, K. Development of yeast cell factories for consolidated bioprocessing of lignocellulose to bioethanol through cell surface engineering. *Biotechnol Adv,* **30,** 1207-1218 (2012).
Uihlein, A. and Schbek L. Environmental impacts of a lignocellulosic feedstock biorefinery system: an assessment. *Biomass Bioenerg.* **33** (2009).
Urban, M. & Adamczak, M. Exploration of metagenomes for new enzymes useful in food biotechnology – A review. *PJSNF*. **58,** 11-22 (2008)
Ventorino, V. *et al*. Response to salinity stress of *Rhizobium leguminosarum* bv. *viciae* strains in the presence of different legume host plants. *Ann. Microbiol*. **62,** 811-823 (2012).
Ventorino, V. *et al*. Exploring the microbiota dynamics related to vegetable biomasses degradation and study of lignocellulose-degrading bacteria for industrial biotechnological application. *Sci. Rep.* **5,** 8161 (2015).
Ventorino, V. *et al*. Chestnut green waste composting for sustainable forest management: Microbiota dynamics and impact on plant disease control. *J. Environ. Manage.* **166,** 168-177 (2016).
Wang, C. *et al.* New insights into the structure and dynamics of actinomycetal community during manure composting. *Appl. Microbiol. Biotechnol.* **98,** 3327-3337 (2014).
Wang, L., Sharifzadeh, M., Templerbd, R. & Murphy R.J. Technology performance and economic feasibility of bioethanol production from various waste papers. *Energ Environ Sci.* **5,** 5717-5730 (2012).
Wang, Y. *et al*. A novel lignin degradation bacterial consortium for efficient pulping. *Bioresour. Technol.* **139,** 113-119 (2013).
Warnecke, F. *et al*. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* **450,** 560-565 (2007).
Wenzel, M., SchoÈnig, I., Berchtold, M., KaÈmpfer, P. & KoÈnig, H. Aerobic and facultatively anaerobic cellulolytic bacteria from the gut of the termite Zootermopsis angusticollis. *J. Appl. Micobiol.* **92,** 32-40 (2002).

Whitehouse, N., Olson, V., Schwab, C., Chesbro, W., Cunningham, K. & Lykos, T. Improved techniques for dissociating particle-associated mixed ruminal microorganisms from ruminal digesta solids. *Journal of Animal Science,* **72,** 1335–1343 (1994).
Whitman, W., Coleman, D. & Wiebe, W. Prokaryotes: The unseen majority. *Proc Natl Acad Sci.* **95,** 6578–6583 (1998).
Woo, H. L., Terry, Hazena, C., Simmonsa, B. A. & DeAngelis, K. M. Enzyme activities of aerobic lignocellulolytic bacteria isolated fromwet tropical forest soils. *Syst. Appl. Microbiol.* **37,** 60-67 (2014).
Wood, T.M., Wilson, C.A. & Stewart, C.S. Preparation of the cellulase from the cellulolytic anaerobic rumen bacterium Ruminococcus albus and its release from the bacterial cell wall. *Biochem J.* **205,** 129–137 (1982).
Wu, F-C., Huang, S-S. & Shih, I-L. Sequential hydrolysis of waste newspaper and bioethanol production from the hydrolysate. *Bioresour Technol.* **167,** 159-168 (2014).
Xu, Q., Adney, W. S., Ding, S. Y. & Michael, H. E., in *Industrial Enzymes. Structure, Function and Applications.* (Springer, 2007), pp. 35-50.
Zhang, X-Z. & Zhang Y-H. P. Cellulases: Characteristics, Sources, Production, and Applications. In: Bioprocessing Technologies in Biorefinery for Sustainable Production of Fuels, Chemicals, and Polymers, First Edition. Edited by Shang-Tian Yang, Hesham A. El-Enshasy, and Nuttha Thongchul (2013).
Zhao, X., Cheng, K. & Liu, D. Organosolv pretreatment of lignocellulosic biomass for enzymatic hydrolysis. *Appl Microbiol Biotechnol.* **82,** 815-27 (2009).

Appendices

Appendix I

List of communications

- Ventorino, V., Faraco, V., **Montella, S.**, Amore, A., Ercolini, D.& Pepe, O. Ecosystem structure and potential functional diversity in microbiota-adapted lignocellulosic biomass. 13th Symposium on Bacterial Genetics and Ecology (BAGECO 2015) – 14/18 June 2015. Milan, Italy.

List of pubblications

- **Montella, S.**, Amore, A. & Faraco, V. Metagenomics for the development of new biocatalysts to advance lignocellulose saccharification for bioeconomic development. *Crit. Rev. Biotechnol*. **18,** 1-12 (2015).
- Giacobbe, S., Balan, V., **Montella, S.**, Fagnano, M., Mori, M. & Faraco, V. Assessment of bacterial and fungal (hemi)cellulose-degrading enzymes in saccharification of ammonia fibre expansion-pretreated Arundo donax. *Appl Microbiol Biotechnol.* (2015).
- **Montella, S.**, Balan, V., da Costa Sousa, L., Gunawan, C., Giacobbe, S., Pepe, O. & Faraco, V. Saccharification of newspaper waste after ammonia fiber expansion or extractive ammonia. *AMB Express*. (2016).

Experience in foreign laboratory

Visiting scholar at **Department of Chemical Engineering and Materials Science, DOE Great Lakes Bioenergy Research Center, Michigan State University**, East Lansing, MI 48823, USA for performing the enzymatic hydrolysis of *Arundo donax L.* and Newspaper Waste after Ammonia Fibre EXpansion (AFEX) and Extractive Ammonia (EA) pretreatment. The research activities were supervised by Prof. Venkatesh Balan. (June/October 2013).

Appendix II

Critical Reviews in Biotechnology

http://informahealthcare.com/bty
ISSN: 0738-8551 (print), 1549-7801 (electronic)

Crit Rev Biotechnol, Early Online: 1–12
 DOI: 10.3109/07388551.2015.1083939

Taylor & Francis
Taylor & Francis Group

REVIEW ARTICLE

Metagenomics for the development of new biocatalysts to advance lignocellulose saccharification for bioeconomic development

Salvatore Montella, Antonella Amore*, and Vincenza Faraco

Department of Chemical Sciences, University of Naples "Federico II", Complesso Universitario Monte S. Angelo, Naples, Italy

Abstract

The world economy is moving toward the use of renewable and nonedible lignocellulosic biomasses as substitutes for fossil sources in order to decrease the environmental impact of manufacturing processes and overcome the conflict with food production. Enzymatic hydrolysis of the feedstock is a key technology for bio-based chemical production, and the identification of novel, less expensive and more efficient biocatalysts is one of the main challenges. As the genomic era has shown that only a few microorganisms can be cultured under standard laboratory conditions, the extraction and analysis of genetic material directly from environmental samples, termed metagenomics, is a promising way to overcome this bottleneck. Two screening methodologies can be used on metagenomic material: the function-driven approach of expression libraries and sequence-driven analysis based on gene homology. Both techniques have been shown to be useful for the discovery of novel biocatalysts for lignocellulose conversion, and they enabled identification of several (hemi)cellulases and accessory enzymes involved in (hemi)cellulose hydrolysis. This review summarizes the latest progress in metagenomics aimed at discovering new enzymes for lignocellulose saccharification.

Keywords

Biocatalysts, biorefinery, cellulases, hemicellulases, metagenome, nonedible biomasses

History

Received 20 February 2015
Revised 23 May 2015
Accepted 21 June 2015
Published online 16 September 2015

Introduction

Production of electricity, fuel and other products, such as solvents and plastics, is highly dependent on fossil sources (Uihlein & Schebek, 2009). The dramatic increase in petroleum prices, the shortage of fossil reserves and increasing concerns regarding environmental impacts, especially those related to greenhouse gas (GHG) emissions and global warming, are encouraging the search for new renewable resources of energy and materials. As an alternative to fossil-based products, several novel strategies have been developed for the conversion of lignocellulose to obtain building blocks and high-value-added biochemicals (Donate, 2014; Faraco & Hadar, 2011; Hara et al., 2014; Liguori et al., 2013; Tuck et al., 2012). Lignocellulose is the most abundant biomass on Earth, and its use for chemical production does not pose problems of conflict with food production. To maximize the competitiveness and minimize waste production and the environmental impact of the production processes, the complete valorization of all components of lignocellulosic feedstock is being pursued within the biorefinery, through the integrated production of building blocks for high-value-added chemicals and biofuels using renewable biomass in place of petroleum (Daoutidis et al., 2014).

However, the main limitations of industrial production of bio-based chemicals via enzymatic catalysis from lignocellulose are related to the enzymes currently employed in biomass conversion, namely, their high cost, instability and low activity under the required operating conditions. To overcome these challenges, the most advanced research on new biocatalysts exploits metagenomics, a word coined by Handelsman et al. (1998) that refers to an advanced method of exploration of environmental genomic DNA (eDNA) from microorganisms whose cultivation cannot be performed under typical laboratory conditions. It is generally accepted that up to 99% of the microbes in the environment cannot be efficiently cultivated in vitro (Amann et al., 1990). Metagenomics overcomes this drawback, exploring complex DNA samples directly extracted from defined natural environments rather than culturing individual microorganisms (Sharma et al., 2008; Ventorino et al., 2015). At present, metagenomic strategies have been employed to discover novel (hemi)cellulases for lignocellulose saccharification that exhibit better traits, e.g. activity at a wide range of temperatures (Lee et al., 2014) and pH (Duan et al., 2010), high resistance under ionic conditions (Ilmberger et al., 2012) and higher catalytic efficiency, thus also offering economical benefits (Adrio & Demain, 2014).

*Present address: National Renewable Energy Laboratory, Biosciences Center , Golden, CO

Address for correspondence: Vincenza Faraco, Department of Chemical Sciences, University of Naples ''Federico II'', Complesso Universitario Monte S. Angelo, via Cintia, 4 80126 Napoli, Italy. Tel: +39 081 674315. Fax: +39 081 674313. E-mail: vfaraco@unina.it

This review summarizes the main recent achievements of metagenomic approaches aimed at the development of new biocatalysts for lignocellulose saccharification, the techniques adopted, their main bottlenecks, the new biocatalysts developed and their industrial potential.

Metagenomics approaches: expression library versus sequencing of metagenomic DNA

The metagenomic approach is pursued by both the function-driven analysis of expression libraries and sequence-driven analysis (Figure 1). The latter can be achieved by either the homology-based screening of libraries or the direct large-scale sequencing of extracted eDNA.

Any metagenomic approach is a multi-step process including the preparation/pretreatment of the samples and the extraction of the nucleic acids. In the pretreatment step, an enrichment of the sample in the target microorganisms can be performed to improve the specificity of the sample's DNA (Hildebrand et al., 2004; Jiao et al., 2006; Lenk et al., 2012), which is often suitable for samples with a mixture of prokaryotic and eukaryotic cells (Piel et al., 2004). As an alternative, non-enrichment of samples can be used to preserve the natural diversity and relative abundance of microbial communities.

There are two common strategies for eDNA extraction (Xing et al., 2012). Direct extraction consists of sample processing by detergents, enzymes and phenol or chloroform to separate the DNA; indirect extraction includes the separation of the microorganisms from the samples followed by lysis extraction. The former approach shows a high DNA recovery rate, but it yields highly impure and small fragments of extracted DNA; the latter approach has a 10- to 100-fold lower recovery rate (Parachin et al., 2010), but it generates larger DNA fragments. Further purification of eDNA is often essential to remove the humic acids before next analyses (Ghazanfar et al., 2010).

The next key step for either function-driven analysis or a homology-based sequence-driven approach is the construction of a metagenomics library. The selection of appropriate vectors and host systems for this library is crucial. Cosmids and fosmids are the most commonly used vectors for metagenomic library preparations (Henne et al., 1999, 2000; Entcheva et al., 2001) because of their stability, their capacity to accept large fragments of eDNA (25–45 kb) and their feasibility of transfection in the most widely used host *Escherichia coli*. Bacterial artificial chromosomes (BACs) can be used when aiming to detect larger gene fragments (up to approximately 200 kb) encoding more complex pathways (Gillespie et al., 2002; Long et al., 2005). Other microbes,

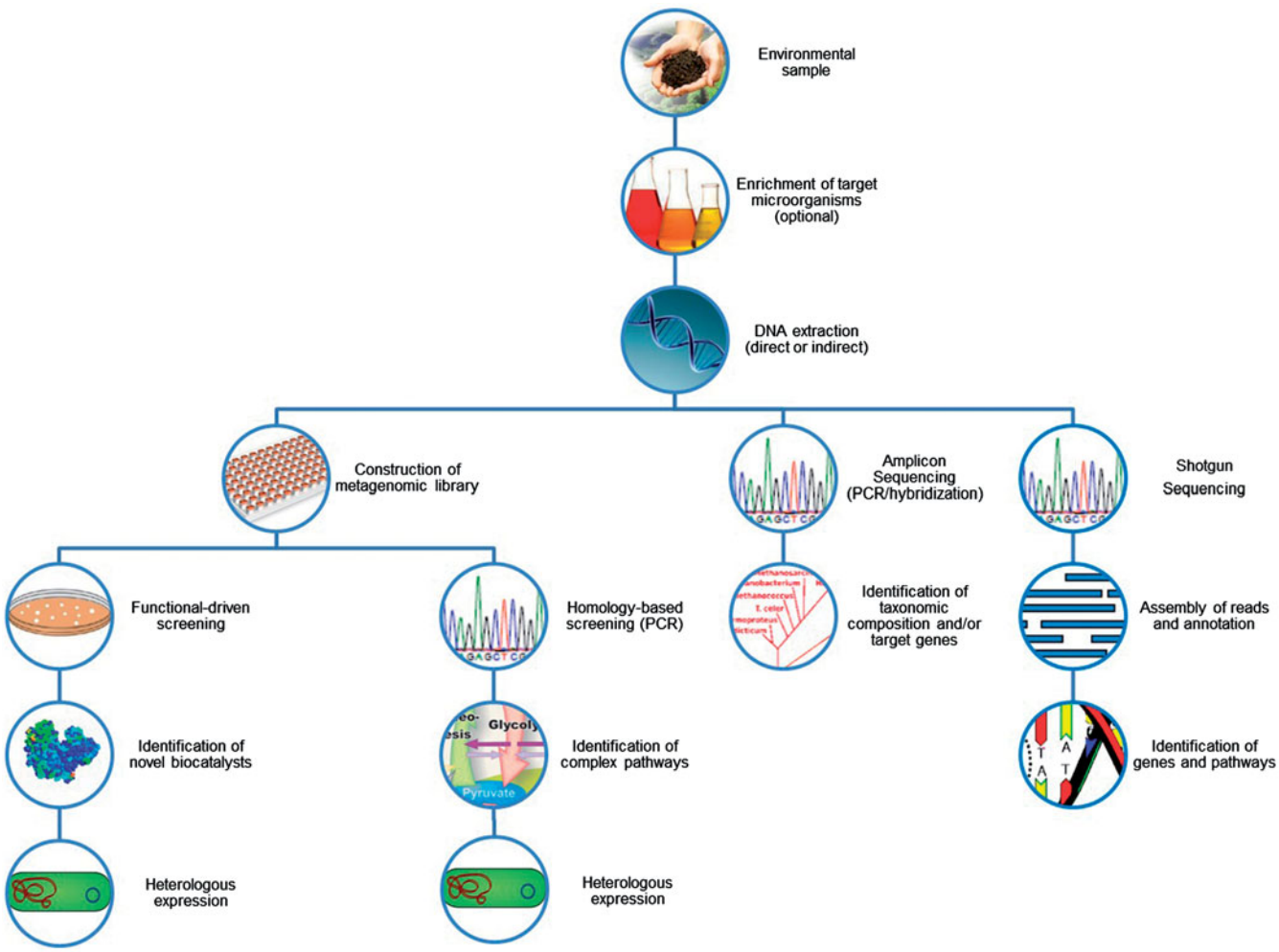


Figure 1. General scheme of metagenomic approaches.

such as *Streptomyces* (Courtois et al., 2003; Wang et al., 2000) and *Pseudomonas* (Wilkinson et al., 2002), have also been employed as hosts to identify novel target genes. However, researchers are developing novel alternative host systems because many eukaryotic genes cannot express functional proteins in these common prokaryotic microorganisms.

The analysis of metagenomic libraries consists of screening the resulting transformant hosts using either of two methods: function-based screening protocols or sequence-driven analysis. The former approach is prevalent for the identification of industrial target proteins. It is a biological activity-based screening that requires the detection of positive clones expressing a desired trait by either direct or indirect detection. Direct detection is achieved by the addition of dyes or chromophore-containing chemicals into the growth medium (Bastien et al., 2013; Rashamuse et al., 2011); the indirect approach consists of the use of host strains able to grow under selective conditions only when expressing the targeted gene product in an active form (Chen et al., 2009). The screening of recombinant positive clones is followed by their sequencing and biochemical analysis. Because functional screening is not performed by sequence similarity to known genes, this type of approach is suitable for identifying novel biocatalysts and biochemical mechanisms.

Alternatively, metagenomic libraries can be screened by homology-based methods using engineered degenerate primers or probes from highly conserved target regions in polymerase chain reaction (PCR) or hybridization to identify complex pathways, such as modular polyketide synthase (PKS) or nonribosomal peptide synthetase (NRPS), that cannot be detected by function-driven analysis (Wilson & Piel, 2013).

The eDNA analysis can also be achieved by directly sequencing all genetic material from a target environment without the creation of a library. This sequence-driven analysis is typically conducted using two approaches: amplicon sequencing and shotgun sequencing. The former is used to filter target gene fragments, such as the 16 S rRNA gene, by PCR amplification of an informative marker and subsequent sequencing. One of the main goals of this type of analysis is the detection of the phylogenetic composition of a microbial environmental population (Klindworth et al., 2013; Soergel et al., 2012). However, if the objective of the metagenomic analysis is to obtain as much information as possible about the community diversity of a specific environment, including the detection of new members and novel genes, a shotgun sequencing strategy must be performed by extracting a large amount of eDNA, followed by random shearing into smaller fragments and direct sequencing (Venter et al., 2004). Large-scale sequencing of eDNA is becoming increasingly common as next-generation sequencing technologies are enabling the sequencing of thousands of genomes in parallel. However, every sequence-driven strategy is based on the analysis of homology to sequences already present in the databases, and thus it is limited to the isolation of new members of already known families or genes that contain highly conserved regions.

Metagenomic for lignocellulose saccharification

Biochemicals from (hemi)cellulose

The hydrolysis of (hemi)cellulose produces monosaccharides and small oligomers, raw materials for a wide range of high-value-added products. Many chemical and biological processes (e.g. gasification, liquefaction, pyrolysis, hydrogenation and fermentation) and/or cascade reactions have been performed to obtain potential industrially relevant chemicals from sugars, such as sorbitol and fructose (used in food and non-food industries), alkyl polyglucosides (applied as surfactant), polylactic acid (mainly used in food packaging) and 5-hydroxymethyl-2-furfural (a versatile molecule involved in the production of solvents, transportation fuels and polymers) (Rasmussen et al., 2014). However, bioethanol is currently the main bulk chemical produced from lignocellulosic biomasses.

However, the processes of lignocellulosic biomass valorization are not competitive as yet, because the biomass conversion is hindered by the recalcitrance of lignocellulose. In fact, the sources of fermentable sugars in lignocellulosic biomass are celluloses with different morphologies and several types of hemicelluloses with complex structures and varying composition. Moreover, the rate and yield of lignocellulose conversion to fermentable sugars remains low due to the physical ligninic barrier surrounding the cellulose and the resistant crystalline structure of cellulose. To develop a competitive process for lignocellulose hydrolysis, the efficient and economic conversion of all polysaccharides must be achieved (Gao et al., 2011; Govindaswamy & Vane, 2007) by developing more efficient biocatalysts.

To achieve this goal, the most advanced studies on the development of new biocatalysts based on cellulases and hemicellulases exploit metagenomics (Table 1) from environmental samples of different habitats, with or without enrichment of the microbial population (Liu et al., 2013; Ni & Tokuda, 2013; Parisutham et al., 2014; Sukharnikov et al., 2011; Xie et al., 2014; Xing et al., 2012).

Cellulases and auxiliary cellulose-degrading enzymes

Cellulases are enzymes that hydrolyse the β (1→4) glycosidic bond. The three main types of cellulases are endoglucanases (EGs) (EC 3.2.1.4), exoglucanases, including cellobiohydrolases (CBHs) (EC 3.2.1.91), and β-glucosidases (BGs) (EC 3.2.1.21).

There is increasing interest in auxiliary enzymes acting towards cellulose by a nonhydrolytic mechanism of depolymerization (Dimarogona et al., 2012). Among these enzymes, lytic polysaccharide monooxygenases (LPMOs) are the most promising class due to their capability of enhancing the efficiency of lignocellulose biomass degradation by acting on polysaccharides that are recalcitrant to cellulases within highly crystalline cellulose (Beeson et al., 2015; Dimarogona et al., 2013).

Industrial cellulose saccharification requires cellulases and auxiliary cellulose-degrading enzymes showing high thermostability, halotolerance and activity under ionic conditions. The main achievements in discovering new EGs, CBHs, BGs

Table 1. Examples of cellulases and hemicellulases recently identified by metagenomic approaches.

Environmental source for metagenomic DNA	Enzyme name	Screening technology*	Reference
Rumen from fistulated, non-lactating Holstein cows	β-glucosidases (SRF2g14, SRF2g18, LAB20g4 and LAB25g2)	F	Del Pozo et al., 2012
Elephant faeces	Putative glycosyl hydrolases (from 82 GH families)	S	Ilmberger et al., 2014
Soil sample covered to pile of cedar wood sawdust or from its enrichment cultures (with Avicel or unbleached kraft pulp)	Endoglucanases (C1A1, C1A2, C2A1, C2A3 and C2P3); xylanase (X2P1)	F	Mori et al., 2014
Rumen fluid from native goat and Saanen hybrid	Lipase GDSL, cellulase, and glycosyl hydrolase family 10	S	Lim et al., 2013
Gut microflora of *Hermetia illucens*	Endoglucanase - glycosyl hydrolase family 5 (CS10)	F	Lee et al., 2014
Gut from the lower termite *Coptotermes gestroi*	Putative cellulases; hemicellulases (endo-1,4-β-xylanases, α-galactosidases, α-N-arabinofuranosidases, xylan 1,4-β-xylosidases, arabinan endo-1,5-α-L-arabinosidases, endo-1,4-β-mannanases, α-glucuronidases, pectinesterases and pectate lyases)	S	Do et al., 2014
Compost microbial communities collected from various farms	Glycosyl hydrolase family 43 (Biof1_09)	F	Sae-Lee & Boonmee, 2014
Bursaphelenchus xylophilus (Bx) samples isolated from pine wilt disease (PWD) epidemic area	Endoglucanase (Cen219)	F	Zhang et al., 2013
Rice straw enriched	Putative glycosyl hydrolases	S	Reddy et al., 2013
Rumen contents of a grass/hay-fed dairy cow	Multifunctional β-glucosidase/β-xylosidase/α-arabinosidase (Bgxa1)	F	Gruninger et al., 2014
Vermicompost prepared with paper mill sludge and dairy fresh sludge	Endoglucanase wild type cmgl504, mutant Cel5A_2R2, engineered Cel5A_2R2-CBM6	F	Telke et al., 2013; Yasir et al., 2013
Fresh slurry from a biogas digester fed with pig ordure and rice straw	Endoglucanase wild type Exo2b and engineered Tc2bh	F	Geng et al., 2012; Yan et al., 2013
Switchgrass-adapted compost	(Hemi)cellulases (including β-glucosidase JMB19063), debranching enzymes, oligosaccharide-degrading enzymes	S	Allgaier et al., 2010; McAndrew et al., 2013
Hindgut microbiota of growing Yorkshire pigs fed a high-fat basal diet supplemented with cellulose	Polygalacturonas, mannanase, cellobiose phosphorylase	F	Wang et al., 2012b
Swamp buffalo (*Bubalus bubalis*) rumen	Glycosyl hydrolase family 5 (Umcel5B29)	S	Cheema et al., 2012
Crop fluid of adult *Achatina fulica* snails	Glycoside hydrolases and carbohydrate-binding modules	S	Cardoso et al., 2012
Bovine rumen	Cellulases (Cel5A and Cel5B)	F	Rashamuse et al., 2013
Rumen from cannulated dairy cow maintained on a grass/hay diet	Endocellulases (including Cel14b22); β-glucosidases; hydrolases active for other substrates	F	Gong et al., 2012
Rice straw compost	Endoglucanase (RSC-EG1)	F	Yi-Fang et al., 2013
Yak rumen	Glycosyl hydrolases of different families	F	Dai et al., 2012
Buffalo rumen contents	Glycosyl hydrolases of different families (included endoglucanase BT-01)	F	Nguyen et al., 2012
Samples from hindguts of a wood-feeding higher termite (*Microcerotermes* sp.)	Cellulase (C0001) and xylanases (X1098.3 and X0012)	F	Thidarat et al., 2012
Equus burchelli faeces	Exoglucanase avicelase (belonging to GH6 superfamily)	S	Chandrasekharaiah et al., 2012a
Grassland soil	Endoglucanase (Cel01); xylanases (Xyn01 and Xyn02)	F	Nacke et al., 2012
Giant pandas faeces	Putative enzymes from 44 different GH families	S	Zhu et al., 2011
Antarctic soil	Cellulases, lipases/esterases, amylases, proteases	F/S	Berlemont et al., 2011
Yak rumen	Bifunctional xylanase/endoglucanase (RuCelA)	F	Chang et al., 2011
Abalone intestine	Endoglucanese (CelAM11)	F	Kim et al., 2011
Microbes adherent to plant fiber incubated in cow rumen	Carbohydrate-active enzymes	S	Hess et al., 2011
Red soil	Endoglucanese (Cel5G)	F	Liu et al., 2013
Buffalo rumen	Endoglucanase (C5614-1)	F	Duan et al., 2010; Duan & Feng, 2010
Green-waste compost added to switchgrass (*Panicum virgatum*)	Putative glycoside hydrolases domains; one active endoglucanase	S	Allgaier et al., 2010

(*continued*)

Table 1. Continued

Environmental source for metagenomic DNA	Enzyme name	Screening technology*	Reference
Subseafloor sediments	β-glucosidases; β-xylosidases endomannanase; endoxylanases,	S	Klippel et al., 2014
Rumen of Jersey cows	Putative Carbohydrate-Active Enzymes	S	Wang et al., 2013
Gut of *Globitermes sulphureus*	β-glucosidase (Bgl-gs1)	F	Wang et al., 2012a
Alkaline polluted soil	β-glucosidase (Bgl1C)	F	Jiang et al., 2011
Soil from sugar refinery	β-glucosidase (Unbgl1A)	F	Lu et al., 2013
Enrichment culture inoculated with an extract of the shipworm *Teredo navalis*; biogas plant; elephant faeces	Cellulases (CelA2, CelA3 and CelA84)	F	Ilmberger et al., 2012
Gut from a longhorned beetle (*Anoplophora glabripennis*)	Putative glycoside hydrolases (cellulases and xylanases); putative laccases, peroxidases and β-etherases; putative LPMOs	S	Scully et al., 2013
Guts and casts of earthworms	Glycosyl hydrolases	F	Beloqui et al., 2010
Baltic Sea water	β-glucosidase (BglMKg)	F	Wierzbicka-Woœ et al., 2013
Compost soil in the vicinity of a hot water spring – author check word please	Xylanase wild type (Mxyl) and mutant (MxylM4)	F	Verma et al., 2013; Verma & Satyanarayana, 2013
Holstein cattle rumen	Xylanase (Xyln-SH1)	F	Cheng et al., 2012
Enriched sugarcane bagasse compost	Endo-β-1,4-xylanase (Xyn3F)	S	Chawannapak et al., 2012
Rumen contents of a grass hay-fed dairy cow	Endo-β-1,4-xylanase (Xyn10N18)	F	Gong et al., 2013
Compost from pig manure and mushroom culture wastes	Xylanase (Xyn10J)	F	Jeong et al., 2012
Rice straw degrading enrichment culture from soil	Xylanase (Umxyn10A)	F	Mo et al., 2010
Sugarcane soil	Endoxylanase (SCXyl)	F	Alvarez et al., 2013
Soil from garden compost	Xylose isomerases (xym1 and xym2)	S/F	Parachin and Gorwa-Grauslund, 2011
Yak (*Bos grunniens*) rumen	β-glucosidase/xylosidase (RuBGX1)	F	Zhou et al., 2011
Rumen of dairy cow	Feruloyl esterase (R_09-02); β-1,4 xylosidase, α-1,5 arabinofur(pyr)anosidase, β-1,4 lactase, α-1,6raffinase, α-1,6stachyase, β-galactosidase and α-1,4glucosidase)	F	Ferrer et al., 2012
Agricultural soil	Type A feruloyl esterase (EstF27)	F	Chen et al., 2012; Sang et al., 2011
Coptotermes formosanus (termite) gut	Feruloyl esterase (Ft3-7)	S	Chandrasekharaiah et al., 2011
Trinervitermes trinervoides (termite) hindgut	Feruloyl esterases (FAE1, FAE2, FAE3, FAE4, FAE5, FAE6, FAE7)	F	Rashamuse et al., 2014
Rusa unicolor and *Equus burchelli* faeces	Feruloyl esterases (Tvms10a and Tvmz2a)	S	Chandrasekharaiah et al., 2012b
Aqueous acidic leachate	Feruloyl esterase (Fae6)	F	Rashamuse et al., 2011
Cellulose-enrichment of gut from earthworm *A. caliginosa*	Type A feruloyl esterases (3A6) and its mutants by random mutagenesis	F	Vieites et al., 2010
China Holstein cow rumen	Feruloyl esterases (FAE-SH1)	F	Cheng et al., 2012
Elephant faeces	Rhamnosidase (RhaB) with a GDSL-like lipase motif an acetyl-xylan esterase motif	F/S	Rabausch et al., 2014
Whole termite abdomens and fungal-comb	β-D-xylosidases and α-L-arabinofuranosidases (GH43 (A3); GH43 (F3); CBM4 (G12); GH51 (G12); GH43 (G12); GH51 (G12))	F	Bastien et al., 2013
Termite gut	Arabinose-specific glycoside hydrolases	F	Arnal et al., 2015
Rumen fluid from a fistulated cow	α-glucoronidase (RUM630-AG)	F	Lee et al., 2012
Forest soil	Lichenase (Mt-lic)	F	Kim et al., 2014
Soil from a cold desert	Endocellulase (CEL8M)	F	Bhat et al., 2013
Forest soil	Pectinase (pg_4)	F	Sathya et al., 2014
Gut from honey bee (*Apis mellifera*)	Carbohydrate breakdown putative enzymes (included pectinases)	S	Engel et al., 2013
Soil from hot springs	Pectinase (PecJKR01His)	S	Singh et al., 2012
Tropical forest soil	Phenol oxidases, peroxidases, β-D-glucosidases, cellobiohydrolases, β-xylopyranosidases, chitinases, CMCases, and xylanases	F	Woo et al., 2014
Canadian soils	Glycoside hydrolases domains (including putative LPMOs)	S	Verastegui et al., 2014
Sugarcane bagasse	Putative (hemi)cellulases, debranching enzymes and various auxiliary activities (including AA9)	S	Mhuantong et al., 2015
Midgut of European Corn Borer *Ostrinia nubilalis*	Putative (hemi)cellulases, debranching enzymes and various Carbohydrate-Binding Modules (including CBM33)	S	Belda et al., 2011

(*continued*)

Table 1. Continued

Environmental source for metagenomic DNA	Enzyme name	Screening technology*	Reference
Anaerobic microbial community decomposing yellow poplar saw dust	Putative glycoside hydrolases, lignin peroxidases and Carbohydrate-Binding Modules (including CBM33)	S	van der Lelie et al., 2012
Rice straw	Putative glycoside hydrolases and Carbohydrate-Binding Modules (including CBM33)	S	Simmons et al., 2014

*F = function-driven analysis; S = sequence-driven analysis.

and LPMOs with these properties by metagenomic approaches are described below and summarized in Table 1.

Endo-1,4-β-glucanases. EGs – endoglucanases – (EC 3.2.1.4) perform the degradation of amorphous cellulose by randomly hydrolysing the internal glycosidic bonds.

The biotechnological potential of novel cellulases from soil metagenomes has been demonstrated. In fact, soils are considered the most diverse microbial habitat on Earth with respect to species diversity and community size. Mori et al. (2014) constructed five metagenomic expression libraries both from raw soil sample which was covered with a pile of cedar wood sawdust and from the same soil enriched with Avicel or unbleached kraft pulp as a carbon source. In this manner, they identified four novel endoglucanases and one novel xylanase, demonstrating the effectiveness of this technique for obtaining cellulases with high activity on crystalline cellulose. Nacke et al. (2012) discovered a highly halotolerant endoglucanase from three large-insert metagenomic libraries produced from different grassland soil samples. Berlemont et al. (2011) identified an endoglucanase showing high thermostability at temperatures above 60 °C while exploring the Antarctic soil metagenome by activity-driven screening.

Compost is another potential source of robust lignocellulolytic enzymes because certain phases of the composting process occur at high temperatures. In DNA extracted from farm compost microbial communities, Sae-Lee & Boonmee (2014) discovered a novel gene encoding a glycosyl hydrolase family 43 exhibiting both endocellulase and endoxylanase activities. Following high-temperature incubation of switchgrass compost, Allgaier et al. (2010) identified one GH9 enzyme among 800 putative glycoside hydrolases detected by sequence-driven analysis. Yasir et al. (2013) prepared a metagenomic library from vermicompost, and the novel cmgl504 endoglucanase was identified and characterized by functional screening. After two rounds of error-prone polymerase chain reaction (PCR) on this endoglucanase and screening of 3000 mutants, Telke et al. (2013) detected amino acid substitutions at various positions in thermotolerant mutants and identified the most heat-tolerant mutant, Cel5A_2R2, which showed a sevenfold increase in thermostability. Furthermore, they fused this mutant with a carbohydrate-binding module (CBM) from *Saccharophagus degradans*, obtaining a protein showing sevenfold higher activity than the native on crystalline cellulose. Yi-Fang et al. (2013) demonstrated the utility of metagenomics in discovering novel cellulolytic genes from the rice-straw compost microbiome. To investigate endoglucanase and xylanase activities, Reddy et al. (2013) constructed a library after high-solid incubation at low and high temperatures (35 and 55 °C, respectively) to enrich the microbial and enzymatic populations that degrade the rice straw. Other endoglucanases were identified by metagenomic approaches from sugarcane bagasse-decomposed soil, cow rumen fluid, and activated sludge from a soda pulp mill, all individually enriched with rice straw (Wongwilaiwalin et al., 2013). The activated sludge from a soda pulp mill was also used to feed a biogas digester in addition to pig ordure (Geng et al., 2012; Yan et al., 2013).

The rumen/gut of various organisms has been revealed as a source of several endoglucanases by metagenomics through both the function-driven (Beloqui et al., 2010; Ilmberger et al., 2012; Kim et al., 2011; Lee et al., 2014; Nguyen et al., 2012; Thidarat et al., 2012; Wang et al., 2012a,b; Zhang et al., 2013) and sequence-driven approaches (Cardoso et al., 2012; Dai et al., 2012; Hess et al., 2011; Lim et al., 2013; Qi et al., 2011; Scully et al., 2013; Zhu et al., 2011). Some of these enzymes are multifunctional, showing both cellulase and xylanase activities (Chang et al., 2011; Ko et al., 2013; Rashamuse et al., 2013). Wang et al. (2013) conducted the first metagenomic study based on a function-driven screening of a fosmid library to study biofilm-colonizing feed particles, where fibre digestion occurs, in the rumen of Jersey cows. Approximately 600 putative glycosyl hydrolases and a new subfamily of GH5 were thus identified. Duan et al. (2010) identified a novel CBM in endoglucanase C5614-1 from the metagenome of buffalo rumen. Ilmberger et al. (2014) investigated the faecal microbiota of a breast- and a plant-fed elephant by metagenome sequencing observing the presence of many enzymes belonging to approximately 80 families. This number is generally twofold higher than the value reported for any of the previously studied samples and still 20% higher with respect to the study on the GH diversity of cow rumen. Rigden et al. (2014) applied bioinformatics methods to assign a function to the GxGYxYP domain, detected in human metagenomes and involved in sugar metabolism. In that study, they suggested that GxGYxYP-containing proteins belonged to a novel family of unknown specificity.

Exo-1,4-β-glucanases. The CBHs - cellobiohydrolases - are exocellulases that cut 2–4 units from the reducing ends – CBH-I (EC 3.2.1.-) – or nonreducing ends – CBH-II (EC

3.2.1.91) – of cellulose. The cooperation between CBHs and EGs is crucial for effectively hydrolysing lignocellulose. Recently, metagenomic sequencing of libraries from different environmental samples has become the most powerful tool for studying uncultured microbes to discover novel CBHs (Berlemont et al., 2011; Cardoso et al., 2012; Hess et al., 2011; Qi et al., 2011; Lim et al., 2013; Scully et al., 2013; Wang et al., 2013; Wongwilaiwalin et al., 2013; Zhu et al., 2011). Chandrasekharaiah et al. (2012a) identified a novel exo-1,4-β-glucanase after the PCR amplification of a gene from the metagenome of an *Equus burchelli* faecal sample. The recombinant protein was capable of hydrolysing both Avicel and carboxymethylcellulose, showing simultaneous exo- and -endocellulosic activity.

β-Glucosidases. The BGs - β-glucosidases - (E.C. 3.2.1.21) perform the degradation of cellobiose and other cellodextrins, playing a key role in enzymatic lignocellulose degradation because of the high cellobiose level produced by EGs and CBHs. Many metagenomes extracted from a variety of environmental samples – including soil (Lu et al., 2013), compost (McAndrew et al., 2013), subseafloor sediments (Klippel et al., 2014) and sea water (Wierzbicka-Woś et al., 2013) – have been analyzed in order to identify novel BGs. Recently, several novel BGs were discovered in different rumen microbiotas (Del Pozo et al., 2012; Gong et al., 2012). Because termites are efficient lignocellulose decomposers, many studies have been conducted to elucidate the termite lignocellulose-degrading systems, with the identification of many cellulose-hydrolyzing enzymes with better characteristics (Ni & Tokuda, 2013), such as activity at various pH values, high glucose tolerance and thermostability (Wang et al., 2012a). Recently, Do et al. (2014) sequenced the metagenome of free-living microbes in the gut of the lower termite *Coptotermes gestroi* using Illumina-based *de novo* sequencing. Approximately 300 ORFs were associated with (hemi)cellulose degrading enzymes – with an abundance of β-glucosidases – and 12 ORFs encoded pectinesterases and pectatelyases.

Lytic polysaccharide monooxygenases. The lytic polysaccharide monooxygenases (LPMOs) are involved in the depolymerization of recalcitrant insoluble polysaccharides from highly crystalline cellulose, catalyzing the cleavage of the internal glycosidic bonds through hydroxylation at the C1, C4 or C6 carbon and subsequent release of oxidized soluble oligosaccharides. This mechanism provides new sites for CBHs and BGs action, enhancing the efficiency of cellulose depolymerization (Beeson et al., 2015). LPMOs contain surface-exposed catalytic copper involved in the regioselective hydroxylation of crystalline cellulose. LPMOs belong to the following carbohydrate-active enzyme (CAZy) families: auxiliary activity family 9 (AA9) and auxiliary activity family 11 (AA11), both formerly reported as glycoside hydrolase family 61 (GH61) and mainly produced by fungi; and auxiliary activity family 10 (AA10), formerly reported as CMB family 33 (CBM33) and mainly produced by chitin-degrading bacteria and viruses (Busk & Lange, 2015, 2013). Only a few of these proteins have been structurally and/or functionally characterized, and even fewer have been discovered by a metagenomic approach. Scully et al. (2013) investigated the wood-degrading system of a gut community from the Asian invasive beetle *Anoplophora glabripennis*. The detected microbial enzymes suggested the presence of all biocatalysts necessary for cellulose degradation, including – albeit in very low abundance – LPMOs (classified as belonging to GH family 61). Verastegui et al. (2014) identified novel putative LPMOs by combining two culture-independent methods: DNA stable-isotope probing (DNA-SIP) and the activity-based functional metagenomic approach. To enrich the microorganisms involved in (hemi)cellulose degradation, three different Canadian soils were first incubated with stable-isotope-labeled carbohydrates (glucose, cellobiose, xylose, arabinose and cellulose). Subsequently, a metagenomic cosmid library was constructed and functionally screened to identify the putative GHs. Recently, Mhuantong et al. (2015) obtained the largest sugarcane bagasse metagenome to date, from which they identified many predicted ORFs matching with CAZymes, including a small number of LPMOs belonging to the AA9 CAZy family. Regarding the CBM33 family (now classified as LPMOs belonging to the AA10 family) discovered by the metagenomic approach, Belda et al. (2011) identified two putative cellulose-degrading enzymes by high-throughput screening of metagenomes from the field-collected European Corn Borer (ECB) midgut. Van der Lelie et al. (2012) found four homologs of CBM33 in a metagenomic library from the microbial community grown on yellow poplar sawdust after enrichment in anaerobic conditions. To develop biocatalysts tolerant to the high temperatures required by the industrial lignocellulose deconstruction process, Simmons et al. (2014) enriched compost-derived microbial communities on rice straw under thermophilic conditions. By metatranscriptomic analysis, they revealed the increased lignocellulose-destructuring ability of the thermophilic community and the over-expression of one gene from *Micromonospora* coding for an enzyme belonging to CBM33.

The discovery of novel LPMOs by metagenomics is a promising research area in the field of biomass conversion due to the proven need to use these biocatalysts in synergism with enzymatic cocktails of cellulases and hemicellulases for improved lignocellulose saccharification (Eibinger et al., 2014).

Hemicellulases

The hemicelluloses are polysaccharides composed of different glyco-units and glycosidic bonds, toward which several enzymes (e.g. xylanase, β-glucanase, mannanase, xyloglucanase, arabinase and esterase) are specific. The main results in the discovery of novel hemicellulases by metagenomic approaches are shown below and summarized in Table 1.

Xylanases (endo-β-xylanases and β-xylosidases). The endo-xylanases – EXs – (E.C. 3.2.1.8) perform the degradation of (glucurono)(arabino)xylan, cleaving the backbone glucosidic bonds in xylan. The β-xylosidases - BXs - (E.C. 3.2.1.37) act on xylooligosaccharides, after the action of EXs toward xylan.

Many novel xylanases have been selected from environmental DNA libraries based on metagenomic analyses

(Cheng et al., 2012a; Jeong et al., 2012). Verma et al. (2013; Verma & Satyanarayana, 2013) expressed an endoxylanase identified from a library constructed close to a hot water spring in *E. coli* BL21 (Fukuoka, Japan): the recombinant xylanase (rMxyl) showed high thermostability and alkali-resistance; furthermore, they improved the thermostability of this novel xylanase by site-directed mutagenesis. Other genes coding for thermotolerant and alkali-tolerant endoxylanases were discovered in metagenomes extracted from microflora in a high-temperature sugarcane bagasse compost (Chawannapak et al., 2012) and in a rice straw-enriched sample (Mo et al., 2010). In sugarcane soil, Alvarez et al. (2013) detected a novel endoxylanase (SCXyl) showing the ability to hydrolyze small xylooligosaccharides. From a metagenomic library constructed from bovine rumen, Gong et al. (2013) retrieved a highly substrate-specific endo-acting xylanase (Xyn10N18) with considerably high pH tolerance.

Several xylanases showing multi-activities were also discovered by metagenomics and could enable the use of single-enzyme resources instead of enzyme cocktails. Zhou et al. (2011) reported the first bifunctional β-glucosidase and xylosidase from the rumen metagenome. Ferrer et al. (2012) identified the first enzyme (R_09-02) that exhibits β-1,4 xylosidase, α-1,5 arabinofur(pyr)anosidase, β-1,4 lactase, α-1,6 raffinase, α-1,6 stachyase, β-galactosidase and α-1,4 glucosidase activities from the calf rumen metagenome.

Glucanase, mannanase, xyloglucan hydrolase and pectinase. β-glucanases (EC 3.2.1.) hydrolyze β(1→3), (1→4), or (1→6) glucan. Kim et al. (2014) identified an enzyme able to efficiently hydrolyze both barley β-glucan and lichenan through function-based screening of a soil metagenomic library (Kim et al., 2014).

Mannanases (EC 3.2.1.78) are involved in the degradation of galacto-glucomannans. A bacterial polysaccharide utilization locus (PUL) is a group of physically connected genes involved in the hydrolysis of a specific glucan. Mackenzie et al. (2015) investigated a PUL from an uncultured *Bacteroidetes* from the rumen population of Arctic reindeer. In addition to two endoglucanases showing activity against β-glucans, xylans and xyloglucan, they identified GH26_i, which hydrolyzed numerous mannans.

Xyloglucan hydrolases (EC 3.2.1.150, 151, 155) belong to the xyloglucan transferase/hydrolase (XTH) family and perform the hydrolysis of xyloglucan. Based on a rumenal microbial metagenomic library, Wong et al. (2010) cloned and expressed a novel exo-glucanase gene (xeg5B) in *E. coli*. The novel enzyme is highly specific toward xyloglucan and oligoxyloglucan.

The degradation of pectin is performed by several pectinolytic enzymes, such as polygalacturonases (EC 3.2.1.15, 67, 82), pectin lyases, pectate lyases (EC 4.2.2.2, 6, 9, 10) and pectin methyl esterases (EC 3.1.1.11). Sathya et al. (2014) constructed a soil metagenomic library from the Western Ghats of India and isolated nine pectinolytic enzymes. From other soil samples from hot springs in northern India, Singh et al. (2012) identified the pectinase PecJKR01 with novel biochemical properties, such as high enzyme activity over a broad pH range (5–9) and a wide temperature range (30–70 °C).

Acetyl xylan esterase and feruloyl esterase. Acetyl and feruloyl ester substituents in hemicelluloses are removed by acetyl xylan esterases (AXEs, EC 3.1.1.72) and feruloyl esterases (FAEs, EC 3.1.1.73), respectively. These enzymes are required for effective exo/endohemicellulase activity.

Developing a metagenomic library from elephant faeces, Rabausch et al. (2014) discovered a new subclass of bacterial B type α-l-rhamnosidase with an acetyl-xylan esterase at the N-terminal region.

Based on function-driven assays, several FAEs have been screened from libraries of different environmental DNAs, such as termite (Chandrasekharaiah et al., 2011) and earthworm gut (Vieites et al., 2010), faeces from *Rusa unicolor* and *Equus burchelli* (Chandrasekharaiah et al., 2012a,b) and leachate (Rashamuse et al., 2011). Recently, Rashamuse et al. (2014) screened a metagenome library prepared from the hindgut of the termite *Trinervitermes trinervoides* to assess feruloyl esterase activities. From this screening, six *fae* genes were successfully heterologously expressed in *E. coli*. From a metagenomic library prepared from China Holstein cow rumen microbes, Cheng et al. (2012b) identified the novel feruloyl esterase FAE-SH1 that showed high thermo-resistance and pH stability. Another FAE with similar properties (EstF27) was characterized by Sang et al. (2011) after functional screening from a soil metagenome.

α-L-Arabinofuranosidase, α-galactosidase and α-glucuronidase. Alpha-L-arabinofuranosidases – AFs – (E.C. 3.2.1.55) remove the arabinose substituent. Arnal et al. (2015) detected a fragment of DNA from termite gut encoding four putative arabinofuranosidases, belonging to families GH43 and GH51. This fragment was revealed to be a polysaccharide utilization locus (PUL) specific for arabinan hydrolysis. Bastien et al. (2013) screened metagenomic libraries from both whole termite abdomens and fungal-comb material to discover hemicellulases. Six novel α-L-arabinofuranosidases and/or β-D-xylosidases were successfully expressed in *E. coli*.

Alpha-galactosidases (E.C. 3.2.1.22) perform the removal of galactose substituents in hemicelluloses, and alpha-glucuronidases (EC 3.2.1.139) are involved in the elimination of glucuronoyl substituents. Lee et al. (2012) generated a metagenomic library from rumen fluid samples collected from a fistulated cow. After screening for α-glucuronidase activity using a high-throughput solid-phase assay, one recombinant enzyme was subcloned into a prokaryotic vector, overexpressed and biochemically characterized.

Bottlenecks and perspectives on metagenomic approaches

Metagenomics is a recent technology, with a number of issues to be resolved.

Regarding function-based screening for the construction of metagenomics libraries, it is advisable to test more suitable expression vectors and host cells to address the issue of the ineffective expression of foreign genes in *E. coli* using the most common vectors. Several studies are being conducted to overcome these limitations, exploring possible innovative and

more robust expression systems, including other hosts (Courtois et al., 2003; Wang et al., 2000; Wilkinson et al., 2002) and innovative engineered vectors (Ferrer et al., 2007; Uchiyama & Miyazaki, 2009). Moreover, functional assays are needed to discover novel biocatalyst genes from metagenomes. The development of high-throughput functional screening methods is essential for reducing the time required in primary screening (Li et al., 2009).

Fungi and other eukaryotic organisms are very important sources of enzymes involved in lignocellulose saccharification. For example, many industrially used cellulases are derived from fungi instead of bacteria. To date, only a few metagenomic cDNA libraries have been constructed (Bailly et al., 2007; Frias-Lopez et al., 2008; Gilbert et al., 2008; Todaka et al., 2007) due to the difficulty in isolating RNA, the instability of mRNA and the difficulty of its separation from other RNA species. However, the large number of novel target biocatalysts identified from metagenomic cDNA libraries (Duan & Feng, 2010) suggests that the construction of libraries from environmental mRNA is a very interesting topic for future research.

Regarding the sequence-driven approach, the complexity of the metagenomes and the limitation of the technology in achieving full coverage of whole-genome sequences of every species are the main issues for the application of metagenomic sequencing to novel biocatalyst discovery. However, DNA sequencing techniques have improved considerably in recent years: from early advances in the traditional Sanger method (Hunkapiller et al., 1991), the technologies have switched to second-generation DNA sequencing. These next-generation platforms – including 454 pyrosequencing (Margulies et al., 2005; Wheeler et al., 2008), Illumina (Turcatti et al., 2008) and SOLiD (Shendure & Ji, 2008) – are undergoing rapid development, thus obtaining enormous amounts of sequenced reads within hours in a more economic and efficient manner. Moreover, several bioinformatics tools are being developed to reassemble and analyze vast metagenomic data rapidly and more accurately (Kunin et al., 2008). The hit rates for the discovery of novel target genes from eDNA show that the development of the shotgun sequencing strategy will be more suitable than function-based screening or amplicon sequencing (Duan & Feng, 2010). Therefore, the combination of all types of metagenomes screening would be useful to move beyond the bottlenecks of each individual method (Warnecke et al., 2007).

Declaration of interest

The authors report no conflicts of interest.

This work was supported by a grant from the Ministero dell'Università e della Ricerca Scientifica – Industrial research project ''*Development of green technologies for production of BIOchemicals and their use in preparation and industrial application of POLImeric materials from agricultural biomasses cultivated in a sustainable way in Campania region (BioPoliS)*'' PON03PE_00107_1, funded in the frame of Operative National Programme Research and Competitiveness 2007-2013 D. D. Prot. N. 713/Ric. 29/10/2010.

References

Adrio JL, Demain AL. (2014). Microbial enzymes: tools for biotechnological processes. Biomolecules, 4, 117–39.

Allgaier M, Reddy A, Park JI, et al. (2010). Targeted discovery of glycoside hydrolases from a switchgrass-adapted compost community. PLoS One, 5, e8812.

Alvarez TM, Goldbeck R, dos Santos C, et al. (2013). Development and biotechnological application of a novel endoxylanase family GH10 identified from sugarcane soil metagenome. PLoS One, 8, e70014.

Amann RI, Binder BJ, Olson RJ, et al. (1990). Combination of 16S rRNA targated oligonucleotide probes with flow-cementry for analysing mixed microbial populations. Appl Environ Microbiol, 56, 1910–25.

Arnal G, Bastien G, Monties N, et al. (2015). Investigating the function of an arabinan utilization locus isolated from a termite gut community. Appl Environ Microbiol, 81, 31–9.

Bailly J, Fraissinet-Tachet L, Verner MC, et al. (2007). Soil eukaryotic functional diversity, a metatranscriptomic approach. ISME J, 1, 632–42.

Bastien G, Arnal G, Bozonnet S, et al. (2013). Mining for hemicellulases in the fungus-growing termite *Pseudacanthotermes militaris* using functional metagenomics. Biotechnol Biofuels, 6, 78–92.

Beeson WT, Vu VV, Span EA, et al. (2015). Cellulose degradation by Polysaccharide Monooxygenases. Annual Review of Biochemistry, 84, 923–46.

Belda E, Padrola L, Peretò J, et al. (2011). Microbial Diversity in teh Mindguts of Field and Lab-Reared Populations of the European Corn Borer Ostrinia nubilalis. PLoS ONE, 6, e21751.

Beloqui A, Nechitaylo TY, Lopez-Cortes N, et al. (2010). Diversity of glycosyl hydrolases from cellulose-depleting communities enriched from casts of two earthworm species. Appl Environ Microbiol, 76, 5934–46.

Berlemont R, Pipers D, Delsaute M, et al. (2011). Exploring the Antarctic soil metagenome as a source of novel cold-adapted enzymes and genetic mobile elements. Rev Argent Microbiol, 43, 94–103.

Bhat A, Riyaz-Ul-Hassan S, Ahmad N, et al. (2013). Isolation of cold-active, acidic endocellulase from Ladakh soil by functional metagenomics. Extremophiles, 17, 229–39.

Busk PK, Lange L. (2013). Function-based classification of carbohydrate-active enzymes by recognition of short, conserved peptide motifs. Applied and Environmental Microbiology, 79, 3380–91.

Busk PK, Lange L. (2015). Classification of fungal and bacterial lytic polysaccharides monooxygenases. BMC Genomics, 16, 368–81.

Cardoso AM, Cavalcante JJV, Maurıcio E, et al. (2012). Metagenomic analysis of the microbiota from the crop of an invasive snail reveals a rich reservoir of novel genes. PLoS One, 7, e48505.

Chandrasekharaiah M, Thulasi A, Bagath M, et al. (2011). Molecular cloning, expression and characterization of a novel feruloyl esterase enzyme from the symbionts of termite (*Coptotermes formosanus*) gut. BMB Rep, 44, 52–7.

Chandrasekharaiah M, Thulasi A, Bagath M, et al. (2012a). Identification of cellulase gene from the metagenome of *Equus burchelli* fecal samples and functional characterization of a novel bifunctional cellulolytic enzyme. Appl Biochem Biotechnol, 167, pp 132–41.

Chandrasekharaiah M, Thulasi A, Vijayarani K, et al. (2012b). Expression and biochemical characterization of two novel feruloyl esterases derived from fecal samples of Rusa unicolor and *Equus burchelli*. Gene, 500, 134–9.

Chang L, Ding M, Bao L, et al. (2011). Characterization of a bifunctional xylanase/endoglucanase from yak rumen microorganisms. Appl Microbiol Biotechnol, 90, 1933–42.

Chawannapak W, Laothanachareon T, Boonyapakron K, et al. (2012). Alkaliphilic endoxylanase from lignocellulolytic microbial consortium metagenome for biobleaching of eucalyptus pulp. J Microbiol Biotechnol, 22, 1636–43.

Cheema TA, Jirajaroenrat K, Sirinarumitr T, Rakshit SK. (2012). Isolation of a gene encoding a cellulolytic enzyme from swamp buffalo rumen metagenomes and its cloning and expression in Escherichia coli. Anim Biotechnol, 23, 261–77.

Chen IC, Lin WD, Hsu SK, et al. (2009). Isolation and characterization of a novel lysine racemase from a soil metagenomic library. Appl Environ Microbiol, 75, 5161–6.

Cheng F, Sheng J, Cai T, et al. (2012b). A protease-insensitive feruloyl esterase from China Holstein cow rumen metagenomic library: expression, characterization, and utilization in ferulic acid release from wheat straw. J Agric Food Chem, 60, 2546–53.
Cheng F, Sheng J, Dong R, et al. (2012a). Novel xylanase from a Holstein Cattle rumen metagenomic library and its application in xylooligosaccharide and ferulic acid production from wheat straw. J Agric Food Chem, 60, 12516–24.
Chen Sk, Wang K, Liub Y, Hu X. (2012). Crystallization and preliminary X-ray analysis of a novel halotolerant feruloyl esterase identified from a soil metagenomic library. Acta Crystallogr Sect F Struct Biol Cryst Commun, 68, 767–70.
Courtois S, Cappellano CM, Ball M, et al. (2003). Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. Appl Environ Microbiol, 69, 49–55.
Dai X, Zhu Y, Luo Y, et al. (2012). Metagenomic insights into the fibrolytic microbiome in yak rumen. PLoS One, 7, e40430.
Daoutidis P, Kelloway A, Marvin WA, et al. (2014). Process systems engineering for biorefineries: new research vistas. Curr Opin Chem Eng, 2, 442–7.
Del Pozo MV, Fernández-Arrojo L, Gil-Martínez J, et al. (2012). Microbial β-glucosidases from cow rumen metagenome enhance the saccharification of lignocellulose in combination with commercial cellulase cocktail. Biotechnol Biofuel, 5, 73–85.
Dimarogona M, Topakas E, Christakopoulos P. (2012). Cellulose degradation by oxidative enzymes. Computational and Structural Biotechnology Journal, 2, e201209015.
Dimarogona M, Topakas E, Christakopoulos P. (2013). Recalcitrant polysaccharide degradation by novel oxidative biocatalysts. Applied Microbiology and Biotechnology, 97, 8455–65.
Do TH, Nguyen TT, Nguyen TN, et al. (2014). Mining biomass-degrading genes through Illumina-based de novo sequencing and metagenomic analysis of free-living bacteria in the gut of the lower termite *Coptotermes gestroi* harvested in Vietnam. J Biosci Bioeng, 118, 665–71.
Donate PM. (2014). Green synthesis from biomass. Chem Biol Technol Agric, 1: 4.
Duan CJ, Feng JX. (2010). Mining metagenomes for novel cellulase genes. Biotechnol Lett, 32, 1765–75.
Duan CJ, Liu JL, Wu X, et al. (2010). Novel carbohydrate-binding module identified in a ruminal metagenomic endoglucanase. Appl Environ Microbiol, 76, 4867–70.
Eibinger M, Ganner T, Bubner P, et al. (2014). Cellulose surface degradation by a lytic polysaccharide monooxygenase and its effect on cellulase hydrolytic efficiency. The Journal of Biological Chemistry, 289, 35929–38.
Engel P, Moran NA. (2013). Functional and evolutionary insights into the simple yet specific gut microbiota of the honey bee from metagenomic analysis. Gut Microbes, 4, 60–5.
Entcheva P, Liebl W, Johann A, et al. (2001). Direct cloning from enrichment cultures a reliable strategy for isolation of complete operons and genes from microbial consortia. Appl Env Microbiol, 67, 89–99.
Faraco V, Hadar Y. (2011). The potential of lignocellulosic ethanol production in the Mediterranean basin. Renew Sustain Energy Rev, 15, 252–66.
Ferrer M, Ghazi A, Beloqui A, et al. (2012). Functional metagenomics unveils a multifunctional glycosyl hydrolase from the family 43 catalysing the breakdown of plant polymers in the calf rumen. PLoS One, 7, e38134.
Ferrer M, Golyshina O, Beloqui A, Golyshin PN. (2007). Mining enzymes from extreme environments. Curr Opin Microbiol, 10, 207–14.
Frias-Lopez Shi, Tyson GW, Coleman ML, et al. (2008). Microbial community gene expression in ocean surface waters. Proc Natl Acad Sci USA, 105, 3805–10.
Gao D, Uppungundla N, Chundawat SPS, et al. (2011). Hemicellulases and auxiliary enzymes for improved conversion of lignocellulosic biomass to monosaccharides. Biotechnol Biofuels, 4, 5.
Geng A, Zou G, Yan X, et al. (2012). Expression and characterization of a novel metagenome-derived cellulase Exo2b and its application to improve cellulase activity in *Trichoderma reesei*. Appl Microbiol Biotechnol, 96, 951–62.
Ghazanfar S, Azim A, Ghazanfar MA, et al. (2010). Metagenomics and its application in soil microbial community studies: biotechnological prospects. J Anim Plant Sci, 6, 611–22.
Gilbert JA, Field D, Huang Y, et al. (2008). Detection of large numbers of novel sequences in the metatranscriptomes of complex marine microbial communities. PLoS One, 3, e3042.
Gillespie DE, Brady SF, Bettermann AD, et al. (2002). Isolation of antibiotics turbomycin a and B from a metagenomic library of soil microbial DNA. Appl Environ Microbiol, 68, 4301–6.
Gong X, Gruninger RJ, Qi M, et al. (2012). Cloning and identification of novel hydrolase genes from a dairy cow rumen metagenomic library and characterization of a cellulase gene. BMC Res Notes, 6, 566–76.
Gong X, Gruniniger RJ, Forster RJ, et al. (2013). Biochemical analysis of a highly specific pH stable xylanase gene identified from a bovine rumen-derived metagenomic library. Appl Microbiol Biotechnol, 97, 2423–31.
Govindaswamy S, Vane LM. (2007). Kinetics of growth and ethanol production on different carbon substrates using genetically enigineered xylose-fermenting yeast. Bioresour Technol, 98, 677–85.
Gruningerg RJ, Gong X, Forster RJ, McAllister TA. (2014). Biochemical and kinetic characterization of the multifunctional β-glucosidase/β-xylosidase/a-arabinosidase, Bgxa1. Appl Microbiol Biotechnol, 98, 3003–12.
Handelsman J, Rondon MR, Brady SF, et al. (1998). Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chem Biol, 5, 245–9.
Hara KY, Araki M, Okai N, et al. (2014). Development of bio-based fine chemical production through synthetic bioengineering. Microb Cell Fact, 13, 173–90.
Henne A, Daniel R, Schmitz RA, Gottschalk G. (1999). Construction of environmental DNA libraries in *Escherichia coli* and screening for the presence of genes conferring utilization of 4-hydroxybutyrate. Appl Environ Microbiol, 65, 3901–7.
Henne A, Scmitz RA, Boemeke M, et al. (2000). Screening of Environmental DNA libraries for the presence of genes conferring lipolytic activity on *Escherichia coli*. Appl Environ Microbiol, 66, 3113–16.
Hess M, Sczyrba A, Egan R, et al. (2011). Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. Science, 331, 463–7.
Hildebrand M, Waggoner LE, Liu H, et al. (2004). bryA: an unusual modular polyketide synthase gene from the uncultivated bacterial symbiont of the marine bryozoan Bugula neritina. Chem Biol, 11, 1543–52.
Hunkapiller T, Kaiser R, Koop B, Hood L. (1991). Large-scale and automated DNA sequence determination. Science, 254, 59–67.
Ilmberger N, Gullert S, Dannenberg J, et al. (2014). A comparative metagenome survey of the fecal microbiota of a breast- and a plant-fed Asian elephant reveals an unexpectedly high diversity of glycoside hydrolase family enzymes. PLoS One, 9, e106707.
Ilmberger N, Meske D, Juergensen J, et al. (2012). Metagenomic cellulases highly tolerant towards the presence of ionic liquids—linking thermostability and halotolerance. Appl Microbiol Biotechnol, 95, 135–46.
Jeong YS, Na HB, Kim SK, et al. (2012). Characterization of xyn10J a novel family 10 xylanase from a compost metagenomic library. Appl Biochem Biotechnol, 166, 1328–39.
Jiao J, Wang H, Zeng Y, Shen Y. (2006). Enrichment for microbes living in association with plant tissues. J Appl Microbiol, 100, 830–7.
Kim D, Kim SN, Baik KS, et al. (2011). Screening and characterization of a cellulase gene from the gut microflora of abalone using metagenomic library. J Microbiol, 49, 141–5.
Kim SY, Oh DB, Kwon O. (2014). Characterization of a lichenase isolated from soil metagenome. J Microbiol Biotechnol, 24, 1699–706.
Klindworth A, Pruesse E, Schweer T, et al. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res, 41, e1.
Klippel B, Sahm K, Basner A, et al. (2014). Carbohydrate-active enzymes identified by metagenomic analysis of deep-sea sediment bacteria. Extremophiles, 18, 853–63.
Ko KC, Lee JH, Han Y, et al. (2013). A novel multifunctional cellulolytic enzyme screened from metagenomic resources representing ruminal bacteria. Biochem Biophys Res Commun, 441, 567–72.

Kunin V, Copeland A, Lapidus A, et al. (2008). A bioinformaticians guide to metagenomics. Microbiol Mol Biol Rev, 72, 557–78.

Lee CC, Kibblewhite RE, Wagschal K, et al. (2012). Isolation of a-glucuronidase enzyme from a rumen metagenomic library. Protein J, 31, 206–11.

Lee CM, Lee YS, Seo SH, et al. (2014). Screening and characterization of a novel cellulase gene from the gut microflora of Hermetia illucens using metagenomic library. J Microbiol Biotechnol, 24, 1196–206.

Lenk S, Moraru C, Hahnke Arnds J, et al. (2012). Roseobacter clade bacteria are abundant in coastal sediments and encode a novel combination of sulfur oxidation genes. ISME J, 6, 2178–87.

Li LL, McCorkle SR, Monchy S, et al. (2009). Bioprospecting metagenomes: glycosyl hydrolases for converting biomass. Biotechnol Biofuels, 2, 10.

Liguori R, Amore A, Faraco V. (2013). Waste valorization by biotechnological conversion into added value products. Appl Microbiol Biotechnol, 97, 6129–47.

Lim S, Seo J, Choi H, et al. (2013). Metagenome analysis of protein domain collocation within cellulase genes of goat rumen microbes. Asian Australas J Anim Sci, 26, 1144–51.

Liu G, Qin Y, Li Z, Qu Y. (2013). Development of highly efficient low-cost lignocellulolytic enzyme systems in the post-genomic era. Biotechnol Adv, 31, 962–75.

Long PF, Dunlap WC, Battershill CN, Jaspars M. (2005). Shotgun cloning and heterologous expression of the patellamide gene cluster as a strategy to achieving sustained metabolite production. Chembiochem, 6, 1760–5.

Lu J, Du L, Wei Y, et al. (2013). Expression and characterization of a novel highly glucose-tolerant beta-glucosidase from a soil metagenome. Acta Biochim Biophys Sin, 45, 664–73.

Mackenzie AK, Naas AE, Kracun S, et al. (2015). A polysaccharide utilization locus from an uncultured bacteroidetes phylotype infers ecological adaptation and substrate versatility. Appl Environ Microbiol, 81, 187–95.

Margulies M, Egholm M, Altman WE, et al. (2005). Genome sequencing in microfabricated high-density picolitre reactors. Nature, 437, 376–80.

McAndrew RP, Park JI, Heins RA, et al. (2013). From soil to structure a novel dimeric beta-glucosidase belonging to glycoside hydrolase family 3 Isolated from compost using metagenomic analysis. J Biol Chem, 288, 14985–92.

Mhuantong W, Charoensawan V, Kanokratana P, et al. (2015). Comparative analysis of sugarcane bagasse metagenome reveals unique and conserved biomass-degrading enzymes among lignocellulolytic microbial communities. Biotechnology for Biofuel, 8, 16–32.

Mo XC, Chen Cl, Pang H, et al. (2010). Identification and characterization of a novel xylanase derived from a rice straw degrading enrichment culture. Appl Microbiol Biotechnol, 87, 2137–46.

Mori T, Kamei I, Hirai H, Kondo R. (2014). Identification of novel glycosyl hydrolases with cellulolytic activity against crystalline cellulose from metagenomic libraries constructed from bacterial enrichment cultures. SpringerPlus, 3, 365–71.

Nacke H, Engelhaupt M, Brady S, et al. (2012). Identification and characterization of novel cellulolytic and hemicellulolytic genes and enzymes derived from German grassland soil metagenomes. Biotechnol Lett, 34, 663–75.

Nguyen NH, Maruset L, Uengwetwanit T, et al. (2012). Identification and characterization of a cellulase-encodic gene from the buffalo rumen metagenomic library. Biosci Biotechnol Biochem, 76, 1075–84.

Ni J, Tokuda G. (2013). Lignocellulose-degrading enzymes from termites and their symbiotic microbiota. Biotechnol Adv, 31, 838–50.

Parachin NS, Gorwa-Grauslund MF. (2011). Isolation of xylose isomerases by sequence- and function-based screening from a soil metagenomic library. Biotechnol Biofuels, 4, 9.

Parachin NS, Schelin J, Norling B, et al. (2010). Flotation as a tool for indirect DNA extraction from soil. Appl Microbiol Biotechnol, 87, 1927–33.

Parisutham V, Kim TH, Lee SK. (2014). Feasibilities of consolidated bioprocessing microbes: from pretreatment to biofuel production. Bioresour Technol, 161, 431–40.

Piel J, Hui D, Wen G, et al. (2004). Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the marine sponge Theonella swinhoei. Proc Natl Acad Sci USA, 101, 16222–7.

Sharma P, Kumari H, Kumar M, et al. (2008). From bacterial genomics to metagenomics: concept tools and recent advances. Indian J Microbiol, 48, 173–94.

Qi M, Wang P, OToole N, et al. (2011). Snapshot of the eukaryotic gene expression in muskoxen rumen—a metatranscriptomic approach. PLoS One, 6, e20521.

Rashamuse K, Ronneburg T, Sanyika W, et al. (2014). Metagenomic mining of feruloyl esterases from termite enteric flora. Appl Microbiol Biotechnol, 98, 727–37.

Rashamuse K, Sanyika W, Ronneburg T, Brad D. (2011). A feruloyl esterase derived from a leachate metagenome library. BMB Rep, 45, 14–19.

Rashamuse KJ, Visser DF, Hennessy F, et al. (2013). Characterisation of Two bifunctional cellulase–xylanase enzymes isolated from a bovine rumen metagenome library. Curr Microbiol, 66, 145–51.

Rasmussen H, Sorensen HR, Meyer AS. (2014). Formation of degradation compounds from lignocellulosic biomass in the biorefinery: sugar reaction mechanisms. Carbohydr Res, 385, 45–57.

Reddy AP, Simmons CW, Khudyakov J, et al. (2013). Discovery of microorganisms and enzymes involved in high-solids decomposition of rice straw using metagenomic analyses. PLoS One, 8, e77985.

Rigden DJ, Eberhardt RY, Gilbert HJ, et al. (2014). Structure- and context-based analysis of the GxGYxYP family reveals a new putative class of glycoside hydrolase. BMC Bioinformatics, 15, 196–207.

Sae-Lee R, Boonmee A. (2014). Newly derived GH43 gene from compost metagenome showing dual xylanase and cellulase activities. Folia Microbiol, 59, 409–17.

Sang S, Li G, Hu X, Liu Y. (2011). Molecular cloning overexpression and characterization of a novel feruloyl esterase from a soil metagenomiclibrary. J Mol Microbiol Biotechnol, 20, 196–203.

Sathya TA, Jacob AM, Khan M. (2014). Cloning and molecular modelling of pectin degrading glycosyl hydrolase of family 28 from soil metagenomic library. Mol Biol Rep, 41, 2645–56.

Scully ED, Geib SM, Hoover K, et al. (2013). Metagenomic profiling reveals lignocellulose degrading system in a microbial community associated with a wood-feeding beetle. PLoS One, 8, e73827.

Shendure J, Ji H. (2008). Next-generation DNA sequencing. Nature Biotechnol, 26, 1135–45.

Simmons CW, Reddy AP, D'haeseleer P, et al. (2014). Metatranscriptomic analysis of lignocellulolytic microbial communities involved in high-solids decomposition of rice straw. Biotechnology for Biofuel, 7, 495–508.

Singh R, Dhawan S, Singh K, Kaur J. (2012). Cloning expression and characterization of a metagenome derived thermoactive/thermostable pectinase. Mol Biol Rep, 39, 8353–61.

Soergel DAW, Dey N, Knight R, Brenner SE. (2012). Selection of primers for optimal taxonomic classification of environmental 16S rRNA gene sequences. ISME J, 6, 1440–4.

Sukharnikov LO, Cantwell BJ, Podar M, Zhulin IB. (2011). Cellulases: ambiguous nonhomologous enzymes in a genomic pespective. Trends Biotechnol, 29, 473–9.

Telke AA, Zhuang N, Ghatge SS, et al. (2013). Engineering of family-5 glycoside hydrolase (cel5a). From an uncultured bacterium for efficient hydrolysis of cellulosic substrates. PLoS One, 8, e65727.

Thidarat N, Thongaram T, Uengwetwanit T, et al. (2012). Metagenomic analysis of novel lignocellulose-degrading enzymes from higher termite guts inhabiting microbes. J Microbiol Biotechnol, 22, 462–9.

Todaka N, Moriya S, Saita K, et al. (2007). Environmental cDNA analysis of the genes involved in lignocellulose digestion in the symbiotic protist community of Reticulitermes speratus. FEMS Microbiol Ecol, 59, 592–9.

Tuck CO, Pèrez E, Horvàth IT, et al. (2012). Valorization of biomass: deriving more value from waste. Science, 337, 695–9.

Turcatti G, Romieu A, Fedurco M, Tairi AP. (2008). A new class of cleavable fluorescent nucleotides: synthesis and optimization as reversible terminators for DNA sequencing by synthesis. Nucleic Acids Res, 36, e25.

Uchiyama T, Miyazaki K. (2009). Functional metagenomics for enzyme discovery: challenges to efficient screening. Curr Opin Biotechnol, 20, 616–22.

Uihlein A, Schebek L. (2009). Environmental impacts of a lignocellulosic feedstock biorefinery system: an assessment. Biomass Bioenergy, 33, 793–802.

Van der Lelie D, Taghavi S, McCorkle SM, et al. (2012). The Metagenome of an Anaerobic Microbial Community Decomposing Popolar Wood Chips. PLos ONE, 7, e36740.
Venter JC, Remington K, Heidelberg JF, et al. (2004). Environmental genome shotgun sequencing of the Sargasso Sea. Science, 304, 66–74.
Ventorino V, Aliberti A, Faraco V, et al. (2015). Exploring the microbiota dynamics related to vegetable biomasses degradation and study of lignocellulose-degrading bacteria for industrial biotechnological application. Sci Rep, 5, 8161–73.
Verastegui Y, Cheng J, Engel K, et al. (2014). Multisustrate Isotope Labeling and Metagenomic Analysis of Active Soil Bacterial Communities. mBio, 5, e01157–14.
Verma D, Kawarabayasi Y, Miyazaki K, Satyanarayana T. (2013). Cloning expression and characteristics of a novel alkalistable and thermostable xylanase encoding gene. (Mxyl). retrieved from compost-soil metagenome. PLoS One, 8, e52459.
Verma D, Satyanarayana T. (2013). Improvement in thermostability of metagenomic GH11 endoxylanase. (Mxyl). by site-directed mutagenesis and its applicability in paper pulp bleaching process. Ind Microbiol Biotechnol, 40, 1373–81.
Vieites JM, Ghazi A, Beloqui A, et al. (2010). Inter-conversion of catalytic abilities in a bifunctional carboxyl/feruloyl-esterase from earthworm gut metagenome. Microb Biotechnol, 3, 48–58.
Wang GY, Graziani E, Waters B, et al. (2000). Novel natural products from soil DNA libraries in a streptomycete host. Org Lett, 2, 2401–4.
Wang L, Hatem A, Catalyurek UV, et al. (2013). Metagenomic insights into the carbohydrate-active enzymes carried by the microorganisms adhering to solid digesta in the rumen of cows. PLoS One, 8, e78507.
Wang Q, Qian C, Zhang XZ, et al. (2012a). Characterization of a novel thermostable beta-glucosidase from a metagenomic library of termite gut. Enzyme Microb Technol, 51, 319–24.
Wang W, Archbold T, Kimber MS, et al. (2012b). The porcine gut microbial metagenomic library for mining novel cellulases established from growing pigs fed cellulose-supplemented high-fat diets. J Anim Sci, 40, 400–2.
Warnecke F, Luginbühl P, Ivanova N, et al. (2007). Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. Nature, 450, 560–5.
Wheeler DA, Srinivasan M, Egholm M, et al. (2008). The complete genome of an individual by massively parallel DNA sequencing. Nature, 452, 872–6.
Wierzbicka-Woś A, Bartasun P, Cieśliński H, Kur J. (2013). Cloning and characterization of a novel cold-active glycoside hydrolase family 1 enzyme with β-glucosidase β-fucosidase and β-galactosidase activities. BMC Biotechnol, 13, 22–33.
Wilkinson DE, Jeanicke T, Cowan DA. (2002). Efficient molecular cloning of environmental DNA from geothermal sediments. Biotechnol Lett, 24, 155–61.
Wilson MC, Piel J. (2013). Metagenomic approaches for exploiting uncultivated bacteria as a resource for novel biosynthetic enzymology. Chem Biol, 20, 636–47.
Wong DDWS, Chan VJ, McCormack AA, Batt SB. (2010). Cloning and characterization of an exo-xylogucanase from rumenal microbial metagenome. Protein Pept Lett, 17, 803–8.
Wongwilaiwalin S, Laothanachareo T, Mhuantong W, et al. (2013). Comparative metagenomic analysis of microcosm structures and lignocellulolytic enzyme systems of symbiotic biomass-degrading consortia. Appl Microbiol Biotechnol, 97, 8941–54.
Woo HL, Hazen TC, Simmons BA, DeAngelis KM. (2014). Enzyme activities of aerobic lignocellulolytic bacteria isolated from wet tropical forest soils. Syst Appl Microbiol, 37, 60–7.
Xie S, Syrenne R, Sun S, Yuan JS. (2014). Exploration of natural biomass utilization systems. (NBUS). for advanced biofuel — from systems biology to synthetic design. Curr Opin Biotechnol, 27, 195–203.
Xing MN, Zhang XZ, Huang H. (2012). Application of metagenomic techniques in mining enzymes from microbial communities for biofuel synthesis. Biotechnol Adv, 30, 920–9.
Yan X, Geng A, Zhang J, et al. (2013). Discovery of (hemi-)cellulase genes in a metagenomic library from a biogas digester using 454 pyrosequencing. Appl Microbiol Biotechnol, 97, 8173–82.
Yasir M, Khan H, Azam SS, et al. (2013). Cloning and functional characterization of endo-β-14-glucanase gene from metagenomic library of vermicompost. J Microbiol, 51, 329–35.
Yi-Fang Y, Chia-yu CS, Hsion-Wen K, et al. (2013). A metagenomic approach for the identification and cloning of an endoglucanase from rice straw compost. Gene, 519, 360–6.
Zhang L, Fan Y, Zheng H, et al. (2013). Isolation and characterization of a novel endoglucanase from a Bursaphelenchus xylophilus metagenomic library. PLoS One, 8, e82437.
Zhou J, Bao L, Chang L, et al. (2011). Beta-xylosidase activity of a GH3 glucosidase/xylosidase from yak rumen metagenome promotes the enzymatic degradation of hemicellulosic xylans. Lett Appl Microbiol, 54, 79–87.
Zhu L, Wua Q, Dai J, et al. (2011). Evidence of cellulose metabolism by the giant panda gut microbiome. Proc Natl Acad Sci U S A, 108, 17714–19.